

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:17:45 ; Search time 19 Seconds
(without alignments)
233.835 Million cell updates/sec

Title: US-09-868-953A-2

Perfect score: 807
Sequence: 1 MFSTKALLAGLISTALGP.....GIQSHVHINCSDPTAYL 151

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents.AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 806 | 99.9 | 151 | 2 | US-08-563-148E-6 |
| 2 | 802 | 99.4 | 151 | 4 | US-09-194-139-1 |
| 3 | 645 | 79.9 | 133 | 1 | US-07-924-753-15 |
| 4 | 421 | 52.2 | 98 | 4 | US-09-194-139-9 |
| 5 | 256 | 31.7 | 54 | 2 | US-08-563-148E-1 |
| 6 | 196 | 24.3 | 41 | 1 | US-07-924-753-1 |
| 7 | 132 | 16.4 | 27 | 1 | US-07-924-753-14 |
| 8 | 95 | 11.8 | 20 | 2 | US-08-563-148E-5 |
| 9 | 94 | 11.6 | 18 | 1 | US-07-924-753-4 |
| 10 | 90 | 11.2 | 19 | 1 | US-07-924-753-3 |
| 11 | 89 | 11.0 | 19 | 1 | US-07-924-753-12 |
| 12 | 86 | 10.7 | 18 | 1 | US-07-924-753-11 |
| 13 | 74 | 9.2 | 702 | 6 | 5240838-15 |
| 14 | 73 | 9.0 | 14 | 1 | US-07-924-753-10 |
| 15 | 70.5 | 8.7 | 625 | 3 | US-08-996-139-15 |
| 16 | 70.5 | 8.7 | 625 | 4 | US-08-995-659-15 |
| 17 | 70.5 | 8.7 | 625 | 4 | US-09-215-649A-15 |
| 18 | 70.5 | 8.7 | 625 | 4 | US-09-577-780-15 |
| 19 | 70 | 8.7 | 10182 | 4 | US-09-134-001C-3159 |
| 20 | 69 | 8.6 | 14 | 1 | US-07-924-753-9 |
| 21 | 67.5 | 8.4 | 2285 | 4 | US-09-308-375-2 |
| 22 | 67 | 8.3 | 14 | 4 | US-09-194-139-10 |
| 23 | 66 | 8.2 | 2004 | 1 | US-08-375-709-15 |
| 24 | 66 | 8.2 | 2004 | 1 | US-08-752-929-15 |
| 25 | 66 | 8.2 | 2004 | 1 | US-09-090-793-9 |
| 26 | 65.5 | 8.1 | 406 | 4 | US-08-293-778-24 |
| 27 | 65.5 | 8.1 | 406 | 1 | US-08-293-778-24 |

| | | | | | | |
|----|------|-----|------|---|------------------|--------------------|
| 28 | 65.5 | 8.1 | 406 | 2 | US-08-955-471-5 | Sequence 5, Appli |
| 29 | 65.5 | 8.1 | 406 | 5 | PCT-US92-10242-5 | Sequence 5, Appli |
| 30 | 65.5 | 8.1 | 444 | 1 | US-08-475-845-2 | Sequence 2, Appli |
| 31 | 65.5 | 8.1 | 444 | 2 | US-08-327-690-2 | Sequence 2, Appli |
| 32 | 65.5 | 8.1 | 444 | 2 | US-08-660-289-2 | Sequence 2, Appli |
| 33 | 65.5 | 8.1 | 444 | 2 | US-08-537-807-2 | Sequence 2, Appli |
| 34 | 65.5 | 8.1 | 444 | 3 | US-08-871-003-2 | Sequence 2, Appli |
| 35 | 65.5 | 8.1 | 444 | 2 | US-08-464-233-2 | Sequence 2, Appli |
| 36 | 65.5 | 8.1 | 444 | 4 | US-09-189-607-2 | Sequence 2, Appli |
| 37 | 65.5 | 8.1 | 444 | 4 | US-09-378-907-2 | Sequence 2, Appli |
| 38 | 65.5 | 8.1 | 444 | 5 | PCT-US94-05779-2 | Sequence 2, Appli |
| 39 | 65.5 | 8.1 | 466 | 1 | US-07-882-202A-4 | Sequence 2, Appli |
| 40 | 65.5 | 8.1 | 466 | 1 | US-08-021-615A-4 | Sequence 4, Appli |
| 41 | 65.5 | 8.1 | 466 | 1 | US-08-321-777-4 | Sequence 4, Appli |
| 42 | 65.5 | 8.1 | 466 | 4 | US-09-009-217-14 | Sequence 14, Appli |
| 43 | 65.5 | 8.1 | 466 | 4 | US-09-009-656-14 | Sequence 14, Appli |
| 44 | 65.5 | 8.1 | 466 | 5 | PCT-US93-04493-4 | Sequence 4, Appli |
| 45 | 65.5 | 8.1 | 1241 | 4 | US-09-040-774-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1
US-08-563-148E-6
; Sequence 6, Application US/08563148E
; Patent No. 5929224
; GENERAL INFORMATION:
; APPLICANT: Kazuo SUZUKI et al.
; TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: 663-2, Shligi
; STREET: Misaki-machi
; CITY: Isumi-gun
; STATE: Chiba-ken
; COUNTRY: JAPAN
; ZIP: 299-45
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS v.5
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/563 148E
; FILING DATE: No. 5929224ember 27, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 293233/1994
; FILING DATE: 28-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single strand
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: intermediate fragment
; ORIGINAL SOURCE:
; ORGANISM: human
; TISSUE TYPE: T-cell derived leukemia cells
; PUBLICATION INFORMATION:
; AUTHORS: Kazuo SUZUKI et al.
; TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 to 151
; US-08-563-148E-6

Query Match 99.9%; Score 806; DB 2; Length 151;
 Best Local Similarity 99.3%; Pred. No. 3.1e-88;
 Matches 150; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSTKALLAGLISALAGPWANICAGKSSNEIRTCDRHGGQYSAORSQRPBGVDILC 60
 DB 1 MFSTKALLAGLISALAGPWANICAGKSSNEIRTCDRHGGQYSAORSQRPBGVDILC 60

QY 61 SAGSTVYAPFTGMIVGQEKPYQKNAINNGVRISGRGFCVAMFYIKPIYKGPICKGEKL 120
 DB 61 SAGSTVYAPFTGMIVGQEKPYQKNAINNGVRISGRGFCVAMFYIKPIYKGPICKGEKL 120

QY 121 GTLLPQKVPYPGIOSHVHIENCSSDPTAYL 151
 DB 121 GTLLPQKVPYPGIOSHVHIENCSSDPTAYL 151

RESULT 2
 US-09-194-139-1
 ; Sequence 1, Application US/09194139A
 ; Patent No. 6306608

GENERAL INFORMATION:

APPLICANT: Atai, Takao
 TITLE OF INVENTION: ANTI-HUMAN LECT2 ANTIBODY, CELLS PRODUCING THE SAME,
 TITLE OF INVENTION: AND METHOD AND KIT FOR ASSAYING THE SAME

FILE REFERENCE: US application 09/194,139
 CURRENT APPLICATION NUMBER: US/09/194,139A

CURRENT FILING DATE: 1998-11-25
 EARLIER APPLICATION NUMBER: JP 8-132160

EARLIER FILING DATE: 1996-05-27
 NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1

LENGTH: 151
 TYPE: PRT

ORGANISM: Homo sapiens
 FEATURE:

OTHER INFORMATION: residue 58 is either Valine or Isoleucine
 US-09-194-139-1

Query Match 99.4%; Score 802; DB 4; Length 151;
 Best Local Similarity 99.3%; Pred. No. 9.4e-88;

Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFSTKALLAGLISALAGPWANICAGKSSNEIRTCDRHGGQYSAORSQRPBGVDILC 60
 DB 1 MFSTKALLAGLISALAGPWANICAGKSSNEIRTCDRHGGQYSAORSQRPBGVDILC 60

QY 61 SAGSTVYAPFTGMIVGQEKPYQKNAINNGVRISGRGFCVAMFYIKPIYKGPICKGEKL 120
 DB 61 SAGSTVYAPFTGMIVGQEKPYQKNAINNGVRISGRGFCVAMFYIKPIYKGPICKGEKL 120

QY 121 GTLLPQKVPYPGIOSHVHIENCSSDPTAYL 151
 DB 121 GTLLPQKVPYPGIOSHVHIENCSSDPTAYL 151

RESULT 3
 US-07-924-753-15

; Sequence 15, Application US/07924753
 ; Patent No. 5270303

GENERAL INFORMATION:

APPLICANT: Fujio SUZUKI et al.
 TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN

NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:

ADDRESS: Wenderoth, Lind & Ponack
 STREET: 805 Fifteenth Street, N.W., #700

CITY: Washington
 STATE: D.C.

COUNTRY: U.S.A.
 ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 500 kb

COMPUTER: IBM Compatible
 OPERATING SYSTEM: MS-DOS

SOFTWARE: Displaywrite

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/924,753

FILING DATE: 19920804
 CLASSIFICATION: 424

PRIOR APPLICATION DATA:
 APPLICATION NUMBER:

FILING DATE:
 ATTORNEY/AGENT INFORMATION:

NAME: Warren M. Cheek, Jr.
 REGISTRATION NUMBER: 33,367

REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-8850
 TELEFAX:

TELEX:

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
 LENGTH: 133 amino acids

TYPE: AMINO ACID
 STRANDEDNESS: single

TOPOLOGY: linear
 MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:
 ORIGINAL SOURCE:

ORGANISM: bovine
 STRAIN:

INDIVIDUAL ISOLATE:
 DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE: fetal cartilage
 CELL TYPE:

CELL LINE:
 ORGANELLE:

IMMEDIATE SOURCE:
 LIBRARY:

CLONE:
 POSITION IN GENOME:

CHROMOSOME/SEGMENT:
 MAP POSITION:

UNITS:
 FEATURE:

NAME/KEY:
 LOCATION:

IDENTIFICATION METHOD:
 OTHER INFORMATION:

PUBLICATION INFORMATION:
 AUTHORS:

TITLE:
 JOURNAL:

VOLUME:
 ISSUE:

PAGES:
 DATE:

DOCUMENT NUMBER:
 FILING DATE:

PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO:

US-07-924-753-15

Query Match 79.9%; Score 645; DB 1; Length 133;
 Best Local Similarity 87.2%; Pred. No. 3.8e-69;

Matches 116; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 19 GPWANICAGKSSNEIRTCDRHGGQYSAORSQRPBGVDILCSAGSTVYAPFTGMIVGQEK 78
 DB 1 GPWANICAGKSSNEIRTCDRHGGQYTAQRNOKLHGVDVILCSDSSTVYAPFTGMIVGQEK 60

TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: N-terminal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-1

Query Match 24.3%; Score 196; DB 1; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.1e-16;
Matches 34; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 19 GPMANICAGKSSNEIRTCDRHGGCGOYSASRSORPHQGVDL 59
|||||
Db 1 GPMATTCAGKSSNEIRTCDRHGGCGOYTAQRNOKLHGGVDVL 41

RESULT 7
US-07-924-753-14
Sequence 14, Application US/07924753
Patent No. 5270303
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wandering, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-14

Query Match 16.4%; Score 132; DB 1; Length 27;
Best Local Similarity 92.6%; Pred. No. 5e-09;
Matches 25; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 111 KGPRIKGERKLTLLPLQKYVPGIQSHV 137
|||
Db 1 KGIKKGERKLTLLPLQKYVPGIQSHI 27

RESULT 8
US-08-563-148E-5
Sequence 5, Application US/08563148E
Patent No. 5929224
GENERAL INFORMATION:
APPLICANT: Kazuo SUZUKI et al.

;; TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
;; NUMBER OF SEQUENCES: 12
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: 663-2, Shiqi
;; STREET: Misaki-machi
;; CITY: Isumi-gun
;; STATE: Chiba-ken
;; COUNTRY: JAPAN
;; ZIP: 299-45
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: MS-DOS v.5
;; SOFTWARE: Word Perfect 5.1
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/563,148E
;; FILING DATE: No. 5929224ember 27, 1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 293233/1994
;; FILING DATE: 28-NOV-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: C. Bruce Hamburg
;; REGISTRATION NUMBER: 22,389
;; REFERENCE/DOCKET NUMBER: F-5230
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212)986-2340
;; TELEFAX: (212)953-7733
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 20 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single strand
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; FRAGMENT TYPE: intermediate fragment
;; ORIGINAL SOURCE:
;; ORGANISM: human
;; TISSUE TYPE: T-cell derived leukemia cells
;; PUBLICATION INFORMATION:
;; AUTHORS: Kazuo SUZUKI et al.
;; TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
;; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 to 20
;; US-08-563-148E-5
;;
Query Match 11.8%; Score 95; DB 2; Length 20;
Best Local Similarity 85.0%; Pred. No. 8.5e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 129 VYPGIQSHVHENCDSPT 148
DB 1 VYPGIQSHIHENCXDLSDPT 20
;;
RESULT 9
US-07-924-753-4
;; Sequence 4, Application US/07924/753
;; Patent No. 5270303
;; GENERAL INFORMATION:
;; APPLICANT: Fujio SUZUKI et al.
;; TITLE OF INVENTION: CHONROMODULIN-II PROTEIN
;; NUMBER OF SEQUENCES: 15
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: Wenderoth, Lind & Ponack
;; STREET: 805 Fifteenth Street, N.W., #700
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: U.S.A.
;; ZIP: 20005
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: MS-DOS
;; SOFTWARE: DisplayWrite

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/924,753
;; FILING DATE: 19920804
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warren M. Cheek, Jr.
;; REGISTRATION NUMBER: 33,367
;; REFERENCE/DOCKET NUMBER:
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-371-8850
;; TELEFAX:
;;
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 18 amino acids
;; TYPE: AMINO ACID
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL:
;; ANTI-SENSE:
;; FRAGMENT TYPE: internal fragment
;; ORIGINAL SOURCE:
;; ORGANISM: bovine
;; STRAIN:
;; INDIVIDUAL ISOLATE:
;; DEVELOPMENTAL STAGE:
;; HAPLOTYPE:
;; TISSUE TYPE: fetal cartilage
;; CELL TYPE:
;; CELL LINE:
;; ORGANELLER:
;; IMMEDIATE SOURCE:
;; LIBRARY:
;; CLONE:
;; POSITION IN GENOME:
;; CHROMOSOME/SEGMENT:
;; MAP POSITION:
;; UNITS:
;; FEATURE:
;; NAME/KEY:
;; LOCATION:
;; IDENTIFICATION METHOD:
;; OTHER INFORMATION:
;; PUBLICATION INFORMATION:
;; AUTHORS:
;; TITLE:
;; JOURNAL:
;; VOLUME:
;; ISSUE:
;; PAGES:
;; DATE:
;; DOCUMENT NUMBER:
;; FILING DATE:
;; PUBLICATION DATE:
;; RELEVANT RESIDUES IN SEQ ID NO: *
;; US-07-924-753-4
;;
Query Match 11.6%; Score 94; DB 1; Length 18;
Best Local Similarity 88.9%; Pred. No. 9.7e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 129 VYPGIQSHVHENCDSPT 146
DB 1 VYPGIQSHIHENCXDLSD 18
;;
RESULT 10
US-07-924-753-3
;; Sequence 3, Application US/07924/753

Patent No. 5270303
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEetical:
ANTI-SENSE:
FRAGMENT TYPE: Internal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:

FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-3
Query Match 11.2%; Score 90; DB 1; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.00031;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 53 HOGVDLCSAGSTVYAPF 70
DB 2 HOGVDVLCDSGVYAPF 19
RESULT 11
US-07-924-753-12
Sequence 12, Application US/07924753
Patent No. 5270303
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEtical:
ANTI-SENSE:
FRAGMENT TYPE: Internal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:

CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-12

Query Match 11.0%; Score 89; DB 1; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00041;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 68 APTGMTVGGEKPYONKNA 86
|||||1:|||||1:1111
Db 1 APTGKINGGEKPYKNKNA 19

RESULT 12
US-07-924-753-11
; Sequence 11, Application US/07924753
; Patent No. 5270303
; GENERAL INFORMATION:
; APPLICANT: Fujio SUZUKI et al.
; TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,753
; FILING DATE: 19920804
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: Single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-11

Query Match 10.7%; Score 86; DB 1; Length 18;
Best Local Similarity 94.1%; Pred. No. 0.00087;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 79 KPYONKNAINNGVRISG 95
|||||1:|||||1:1111
Db 1 KPYKNKNAINNGVRISG 17

RESULT 13
5240838-15
; Patent No. 5240838
; APPLICANT: LEDBOER, ADRIANUS M.; MAAT, JAN; VERRIPS, CORNELIS
; T.; VISSER, CHRISTIAAN; JANOWICZ, ZBIGNIEW A.; HOLLEMBERG, CORNELIS P.
; TITLE OF INVENTION: REGULATORY SEQUENCES OF ALCOHOL OXIDASE
; (MOX) AND DIHYDROXYACETONESYNTASE (DAS) OF HANSENULA POLYMORPHA
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/587,555
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 300,211
; FILING DATE: 23-JAN-1989
; APPLICATION NUMBER: 759,315
; FILING DATE: 26-JUL-1985
; SEQ ID NO: 15:
; LENGTH: 702
; 5240838-15

Query Match 9.2%; Score 74; DB 6; Length 702;
Best Local Similarity 24.1%; Pred. No. 3.1;
Matches 33; Conservative 22; Mismatches 48; Indels 34; Gaps 7;

OY 11 GLISTALAGPWANICAGKSSNEIRTCDRHCGGYSAORSORPHOGVDILCSAGSTVYAPF 70

Db 525 GLEOKAVELPSSILS-LSRNEVL-----QYLASRQRRRNAGYILIEDAENAEVQI 575
Oy 71 TGMIVGQEKYONKNAIINNGVIRISGRF-----CYKMFYIKPLKYKPT--KKGKEL 120
Db 576 IG--VGAEMFEADKA-----KILGRKFRTRVLSIPCRFLPDEOSIGYRSRVLARKDGRV 628
Oy 121 GTLLPLQKVPYPGIOSHV 137
Db 629 PTIV-----VDARV 637

RESULT 14
US-07-924-753-10
Sequence 10, Application US/07924753
Patent No. 5270303
GENERAL INFORMATION:
APPLICANT: FUJIO SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ. ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:

UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-10

Query Match
Best Local Similarity 9.0%; Score 73; DB 1; Length 14;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 119 KGTLPLQKVPYPG 132
Db 1 KGTLPLQKVPYPG 14

RESULT 15
US-08-996-139-15
Sequence 15, Application US/08996139
Patent No. 6017729
GENERAL INFORMATION:
APPLICANT: Anderson, Dirk M.
APPLICANT: Galibert, Laurent
TITLE OF INVENTION: Receptor Activator of NF-kappaB
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,139
FILING DATE: 22 DECEMBER 1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/064,671
FILING DATE: 14 OCTOBER 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/813,509
FILING DATE: 07 MARCH 1997
APPLICATION DATA:
APPLICATION NUMBER: USSN 08/772,330
FILING DATE: 23 DECEMBER 1996
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2851-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
INFORMATION FOR SEQ. ID NO: 15:
SEQUENCE CHARACTERISTICS:

; LENGTH: 625 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-996-139-15

Query Match 8.7%; Score 70.5; DB 3; Length 625;
Best Local Similarity 21.4%; Pred. No. 7;
Matches 31; Conservative 12; Mismatches 39; Indels 63; Gaps 5;

QY 16 ALAGPWANI-----CAGKSNFIRTCDRHGGCGOYSAQRSGORPHQGVNLCSA 62
DB 309 AAGGPWAEVRDSDRTFLVSEVETQGLSKRTPEDEY-----TDRSQPSTGSLLTQQ 362
QY 63 GSTVYAPFTGMIVGOEKPYQNNAINNGVYISGRGFCVKMFYIKPIKKGPPIKKGKELGT 122
DB 363 GSKSIPPF-----QE-----PLEVGENDS- 381
QY 123 LLPLQKVPYGIQSHVHIENCSSDP 147
DB 382 --LSOCFTGTSTVDSEGCDFTEP 403

Search completed: March 25, 2003, 17:23:00
Job time : 19 secs

RESULT 2
US-10-157-457-9
; Sequence 9, Application US/10157457

```

: Patent No. US20020150568A1
: GENERAL INFORMATION:
: APPLICANT: Williams, John P.
: APPLICANT: McDonald, Jay M.
: APPLICANT: McKenna, Margaret A.
: TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
: FILE REFERENCE: D6294CIP
: CURRENT APPLICATION NUMBER: US/10/157,457
: CURRENT FILING DATE: 2002-05-29
: PRIOR APPLICATION NUMBER: US 09/864,570
: PRIOR FILING DATE: 2001-06-19
: NUMBER OF SEQ ID NOS: 11
: SEQ ID NO 9
: LENGTH: 113
: TYPE: PRT
: ORGANISM: unknown
: FEATURE:
: LOCATION: 38..165
: OTHER INFORMATION: amino acid sequence of region of Lact2
: OTHER INFORMATION: protein homologous to repeat sequences of N
: OTHER INFORMATION: and C terminal ends of the m1m-1 protein
: US-10-157-457-9

```

| | | | | |
|-----------------------|-----------------|--------------|----------|------------|
| Query Match | 87.4% | Score 544 | DB 12 | Length 113 |
| Best Local Similarity | 85.8% | Pred. No. 6 | 3e-54 | |
| Matches 97 | Conservative 10 | Mismatches 6 | Indels 0 | Gaps 0 |

```

Qy      39 HGGCGYSAQRSORPHQGYDILCSAGSTVYAFPTGMYGQEKFPQNKNAIINNGVRLSGRGF  98
          :|||||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 YGCGYSAQRGRHHPGVYDVLCSGSSVYVAFPTGKITVGQEKPYRNKNAINDGIRLSGRGF  600

```

QY 99 CVMFYIKPIKYGPIKKGELGLTLLPQKVYPGIQSHVHIENCDDSDPTAYL 151
||:||||| : |||||:|||||:
Db 61 CVRIIFYIKPIKKGSIKKGEKGLTLPLQKIYPGIQSHVHENCDDSDPTAYL 113

RESULT 3
US-09-884-570-8

```

? sequence B, application US/09864370
? Patent No. US20020114779A1
?
? GENERAL INFORMATION:
? APPLICANT: Williams, John P.
? APPLICANT: McDonald, Jay M.
? APPLICANT: McKenna, Margaret A.
? TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
? FILE REFERENCE: D6294
? CURRENT APPLICATION NUMBER: US/09/884,570
? CURRENT FILING DATE: 2001-06-19
? PRIOR APPLICATION NUMBER: US 60/212,271
? PRIOR FILING DATE: 2000-06-19
? NUMBER OF SEQ ID NOS: 11
?
? SEQ ID NO 8
? LENGTH: 326
? TYPE: prt
? ORGANISM: Gallus gallus
? FEATURE:
? OTHER INFORMATION: amino acid sequence of mim-1 protein
? US-09-884-570-8

```

| | | | | |
|--------------------------|--------|--------------------|------------|-------------|
| Query Match | 53.3%; | Score 430; | DB 10; | Length 326; |
| Best Local Similarity | 50.3%; | Pred. No. 1.6e-40; | | |
| Matches 80; Conservative | 26; | Mismatches 39; | Indels 14; | Gaps 2 |

```

QY      7 LLAGLISTALAGP-----WANICAGKSSMEIRTCDRHGGCGYSAQRSP--P 52
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      7 IALLSLVSTAFARQMEVHPPOGGRHWAQICSGNPFNRIRGCDRGCGGNYASRQKGEK 66

```

Oy 53 HQGV DILCSAGSTVYAPFTGMLVGQEKPYQNKNALNNGVRISRGFCVMFMYIKPLTKYG 112
::
Db 67 HKGV DVLCTDGSIVAPFWQLSGPIREFHNGNALIDGVQLSSGYCVKLVLCIHPRLRYHG 120

QY 113 PIKKGELGTLPLQKVYPGIQSHVHIENCSSDPTAYL 151
|:::| |:::|:::|:::|:::|:::|:::|

Db 127 QIQKGQQLGRMLPMQKVFPGIVSHIHVENCDSPTLL 165

RESULT 4
US-10-157-457-8

Patent No. US20020150568A1

;; GENERAL INFORMATION:
; APPLICANT: Williams, John P.

APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret

TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
;
; REFERENCE: D6394C1P

CURRENT APPLICATION NUMBER: US/10/157,457

PRIOR APPLICATION NUMBER: US 09/884,570

NUMBER OF SEQ ID NOS: 11

; SEQ ID NO 8
; LENGTH: 326

```

; TYPE: PRT
OBTAINING: Cellucell 200

```

FEATURE: 3

OTHER INFO
US-10-157-457-8

| | | | | |
|--------------------------|-------|--------------------|------------|-------------|
| Query Match | 53.3% | Score 430; | DB 12; | Length 326; |
| Best Local Similarity | 50.3% | Pred. No. 1.6e+40; | | |
| Matches 80; Conservative | 26; | Mismatches 39; | Indels 14; | Gaps 2 |

```

QY      7 LLLAGHISTALAGP-----WANICAGKSSNEIRPCDRHGGCGYSAPRSOR--P 52
      : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      7 IALLSLVSTAFARQWEVHPDQOGRHMAQICSGNPFENKIRGCDRYGCGGNYGASNQKGK 66

```

QY HQGVDLICSGSTVYAPFTGMIVGQEKBYQNKNAINNGVRISRGFCYKMFYIKPIKYKG 112
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 67 HKGVADVICTDGSIIYAPFWMQLSGPIRFPHNGNALDDGVQLSSGSYCACLVCIHPIRIRYHG 126

```

QY      113 PIKKGKLTLLPLOKVYPGIQSHVHIENCSDSDPAYLV 151
      114 |:::|:::|:::|:::|:::|:::|:::|
Db      127 QIQGGQQLRMLPMKVPEPGIVSHIYENCQSDPFTLL 165

```

RESULT 5
US-09-884-570-11

```

; sequence 11, Application US/05088570
; Patent No. US20020114779A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Williams, John P

```

APPLICANT: McDonald, Jay M.

TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof

FILE REFERENCE: D0294
CURRENT APPLICATION NUMBER: US/09/884,570

PRIOR APPLICATION NUMBER: US 60/212,271

PRIOR FILING DATE: 2000-
 ;

NO. 11
SEQ ID NO 11

```

; LENGTH: 115
; TYPE: PRT
;

```

ORGANISM: *Gallus gallus*

LOCATION: 200..315

OTHER INFORMATION: terminal end of mim-1 protein homologous

OTHER INFORMATION: to detect protein

Overview Match 43 68% Score 352 DB 10: Length 115:

Best Local Similarity 54.4%; Pred. No. 2.6e-32;
Matches 62; Conservative 25; Mismatches 25; Indels 2; Gaps 1.

; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38651
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004763.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8
; OTHER INFORMATION: SWISSPROT HIT: O14960, EVALU8 8.00e-17
US-09-864-761-38651

Query Match 36.4%; Score 294; DB 10; Length 54;
Best Local Similarity 100.0%; Pred. No. 3.4e-26;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 FCVKMEYIKPIKYGKIGKIGTLPLQKVPYGIOSHVAHIENCDSDPTAYL 151
Db 1 FCVKMEYIKPIKYGKIGKIGTLPLQKVPYGIOSHVAHIENCDSDPTAYL 54

RESULT 10

US-09-884-570-7
; Sequence 7, Application US/09884570
; Patent No. US20020114779A1
; GENERAL INFORMATION:
; APPLICANT: Williams, John P.
; APPLICANT: McDonald, Jay M.
; APPLICANT: McKenna, Margaret A.
; TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
; FILE REFERENCE: D6294
; CURRENT APPLICATION NUMBER: US/09/884,570
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/212,271
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 11

; SEQ ID NO 7
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Gallus gallus
; FEATURE:
; LOCATION: 219..243
; OTHER INFORMATION: amino acid sequence tryptic peptide from
; OTHER INFORMATION: mim-1 protein
US-09-884-570-7

Query Match 10.2%; Score 82; DB 10; Length 25;
Best Local Similarity 56.0%; Pred. No. 0.0093;
Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 55 GVDILCSAGSTYVAPFTGMIVGOEK 79
Db 1 GVDIVCAGATVYAPFSGELSGPVK 25

RESULT 11
US-10-157-457-7
; Sequence 7, Application US/10157457
; Patent No. US20020150568A1
; GENERAL INFORMATION:
; APPLICANT: Williams, John P.
; APPLICANT: McDonald, Jay M.
; APPLICANT: McKenna, Margaret A.
; TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
; FILE REFERENCE: D6294CIP
; CURRENT APPLICATION NUMBER: US/10/157,457
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: US 09/884,570
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 11
; SEQ ID NO 7
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Gallus gallus
; FEATURE:
; LOCATION: 219..243
; OTHER INFORMATION: amino acid sequence tryptic peptide from
; OTHER INFORMATION: mim-1 protein
US-10-157-457-7

Query Match 10.2%; Score 82; DB 12; Length 25;
Best Local Similarity 56.0%; Pred. No. 0.0093;
Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 55 GVDILCSAGSTYVAPFTGMIVGOEK 79
Db 1 GVDIVCAGATVYAPFSGELSGPVK 25

RESULT 12

US-09-884-570-5
; Sequence 5, Application US/09884570
; Patent No. US20020114779A1
; GENERAL INFORMATION:
; APPLICANT: Williams, John P.
; APPLICANT: McDonald, Jay M.
; APPLICANT: McKenna, Margaret A.
; TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
; FILE REFERENCE: D6294
; CURRENT APPLICATION NUMBER: US/09/884,570
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/212,271
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 11
; SEQ ID NO 5
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Gallus gallus
; FEATURE:

LOCATION: 94..115
OTHER INFORMATION: amino acid sequence tryptic peptide from
OTHER INFORMATION: mim-1 protein
US-09-884-570-5

Query Match
Best Local Similarity 61.9%; Pred. No. 0.063;
Matches 13; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 81 YONKNAINGVRISGRFCVK 101
DB 2 FHNGNAIDGVQISGSYCVK 22

RESULT 13
US-10-157-457-5
Sequence 5, Application US/10157457
Patent No. US20020150568A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294CIP
CURRENT APPLICATION NUMBER: US/10/157,457
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: US 09/884,570
PRIOR FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 5
LENGTH: 22
TYPE: PRT
ORGANISM: Gallus gallus
FEATURE:
LOCATION: 94..115
OTHER INFORMATION: amino acid sequence tryptic peptide from
OTHER INFORMATION: mim-1 protein
US-10-157-457-5

Query Match
Best Local Similarity 9.2%; Score 74; DB 12; Length 22;
Matches 13; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 81 YONKNAINGVRISGRFCVK 101
DB 2 FHNGNAIDGVQISGSYCVK 22

RESULT 14
US-09-884-570-6
Sequence 6, Application US/09884570
Patent No. US20020114779A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294
CURRENT APPLICATION NUMBER: US/09/884,570
CURRENT FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/212,271
PRIOR FILING DATE: 2000-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 6
LENGTH: 16
TYPE: PRT
ORGANISM: Gallus gallus
FEATURE:
LOCATION: 293..308
OTHER INFORMATION: amino acid sequence tryptic peptide from
OTHER INFORMATION: mim-1 protein
US-09-884-570-6

Query Match
Best Local Similarity 9.0%; Score 73; DB 10; Length 16;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 129 VYPGIOSHVHENC D 143
DB 1 VFPGLISHVHENC D 15

RESULT 15
US-10-157-457-6
Sequence 6, Application US/10157457
Patent No. US20020150568A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294CIP
CURRENT APPLICATION NUMBER: US/10/157,457
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: US 09/884,570
PRIOR FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 6
LENGTH: 16
TYPE: PRT
ORGANISM: Gallus gallus
FEATURE:
LOCATION: 293..308
OTHER INFORMATION: amino acid sequence tryptic peptide from
OTHER INFORMATION: mim-1 protein
US-10-157-457-6

Query Match
Best Local Similarity 9.0%; Score 73; DB 12; Length 16;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 129 VYPGIOSHVHENC D 143
DB 1 VFPGLISHVHENC D 15

Search completed: March 25, 2003, 17:29:45
Job time : 13 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:16:59 ; Search time 28.5 seconds
(without alignments)
509.344 Million cell updates/sec

Title: US-09-868-953A-2

Perfect score: 807

Sequence: 1 MFSTKALLIAGLISTALGP.....GIQSHVHIENCDSDPTAYL 151

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 645 | 79.9 | 133 | 2 | JH0270 |
| 2 | 433 | 53.7 | 326 | 2 | A33755 |
| 3 | 103.5 | 12.8 | 305 | 2 | T16567 |
| 4 | 80.5 | 10.0 | 417 | 2 | C82840 |
| 5 | 79.9 | 9.8 | 750 | 2 | F90571 |
| 6 | 77.5 | 9.6 | 475 | 2 | F64151 |
| 7 | 76 | 9.4 | 651 | 2 | T42644 |
| 8 | 75 | 9.3 | 240 | 2 | H84197 |
| 9 | 75 | 9.3 | 444 | 2 | T05614 |
| 10 | 74.5 | 9.2 | 268 | 2 | D97054 |
| 11 | 74.5 | 9.2 | 295 | 2 | AC2451 |
| 12 | 74 | 9.2 | 284 | 2 | E97836 |
| 13 | 73 | 9.0 | 350 | 2 | AD0466 |
| 14 | 73 | 9.0 | 710 | 1 | XJHOFK |
| 15 | 73 | 9.0 | 1331 | 1 | XORTDH |
| 16 | 72.5 | 9.0 | 355 | 2 | T51414 |
| 17 | 72 | 8.9 | 1434 | 2 | T32647 |
| 18 | 70.5 | 8.7 | 428 | 2 | T36565 |
| 19 | 70.5 | 8.7 | 454 | 2 | AF3537 |
| 20 | 70.5 | 8.7 | 706 | 2 | T01351 |
| 21 | 70 | 8.7 | 689 | 2 | G84447 |
| 22 | 69 | 8.6 | 310 | 2 | A60967 |
| 23 | 69 | 8.6 | 1347 | 2 | T45632 |
| 24 | 68.5 | 8.5 | 160 | 2 | H82207 |
| 25 | 68.5 | 8.5 | 233 | 2 | AD3344 |
| 26 | 68.5 | 8.5 | 273 | 2 | E81330 |
| 27 | 68.5 | 8.5 | 357 | 2 | A97290 |
| 28 | 68.5 | 8.5 | 608 | 2 | AB3562 |
| 29 | 68 | 8.4 | 184 | 2 | T15360 |

| | | | | | | |
|----|------|-----|------|---|--------|---------------------|
| 30 | 68 | 8.4 | 407 | 2 | G84783 | probable pectinest |
| 31 | 68 | 8.4 | 762 | 2 | D70407 | DNA mismatch repai |
| 32 | 67.5 | 8.4 | 312 | 2 | AE3215 | hypothetical prote |
| 33 | 67.5 | 8.4 | 443 | 2 | I46932 | coagulation factor |
| 34 | 67.5 | 8.4 | 535 | 2 | C36811 | hypothetical prote |
| 35 | 67.5 | 8.4 | 535 | 2 | A37994 | Rf1 protein - saim |
| 36 | 67.5 | 8.4 | 2285 | 2 | T13796 | conserved transglyc |
| 37 | 67 | 8.3 | 271 | 2 | G72380 | probable hypothet |
| 38 | 67 | 8.3 | 609 | 2 | G87496 | peptidase, M23/M37 |
| 39 | 66.5 | 8.2 | 1012 | 2 | T41940 | DNA polymerase - h |
| 40 | 66 | 8.2 | 271 | 2 | A10532 | probable lipoprote |
| 41 | 66 | 8.2 | 335 | 2 | F82016 | probable transposa |
| 42 | 66 | 8.2 | 398 | 2 | A35281 | integumentary muc1 |
| 43 | 66 | 8.2 | 426 | 2 | A81904 | hypothetical prote |
| 44 | 66 | 8.2 | 447 | 2 | B83563 | conserved hypothet |
| 45 | 66 | 8.2 | 554 | 2 | C98133 | component of type |

ALIGNMENTS

RESULT 1

JH0270 Chondromodulin II - bovine

N/Alternate names: cartilage-derived factor

C/Species: Bos primigenius taurus (cattle)

C/Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 17-Mar-1999

C/Accession: JH0270

R/Hiraki, Y.; Inoue, H.; Kondo, J.; Kamizono, A.; Yoshitake, Y.; Shukunami, C.;

J. Biol. Chem. 271, 22657-22662, 1996

A/Title: A novel growth-promoting factor derived from fetal bovine cartilage, c

A/Reference number: JH0270; MUID:96394331; PMID:8798437

A/Accession: JH0270

A/Molecule type: protein

A/Residues: 1-133 <HR>

A/Experimental source: cartilage cell

C/Comment: This protein stimulates proteoglycan synthesis in the growth plate c

Query Match 79.9% Score 645; DB 2; Length 133;

Best local similarity 87.2% Pred. No. 7.3e-58;

Matches 116; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| QY | 19 | GPWANICAGKSSNEIRTCDRHGGQYSAORSORPHQGVILCSASTYVAPFTGMVGOE | 78 |
| DB | 1 | GPWATTCAGKSSNEIRTCDHGGQYTAQNNQKLHGQVDVLCSDGSTVYVAPFTGKIMGOE | 60 |
| QY | 79 | KPYQKNAINNGVRISGRFCYKMFYIKPKIKGPIKKEKLGTLPLQKYYPGIIQSHVH | 138 |
| DB | 61 | KPYKNKNAIINNGVRISGGGFCIKMFYIKPKIKGSIKKEKLGTLPLQKYYPGIIQSHIH | 120 |
| QY | 139 | IENCDSPTAYL 151 | |
| DB | 121 | IENCDSPTAYL 133 | |

RESULT 2

A33755 myb-induced myeloid protein 1 (mim-1) precursor - chicken

N/Alternate names: arginine-specific ADP-ribosyltransferase endogenous substrat

C/Species: Gallus gallus (chicken)

C/Date: 30-Mar-1990 #sequence_revision 30-Mar-1990 #text_change 04-Sep-1998

C/Accession: A33755; S29229

R/Ness, S.A.; Marknell, A.; Graf, T.

Cell 59, 1115-1125, 1989

A/Title: The v-myb oncogene product binds to and activates the promyelocyte-spe

A/Reference number: A33755; MUID:90090611; PMID:2688896

A/Accession: A33755

A/Molecule type: mRNA

A/Residues: 1-326 <NES>

A/Cross-references: GB:M29449; NID:9212341; PID:9212342

R/Tamada, K.; Tsuchiya, M.; Mishima, K.; Shimoyama, M.

FEBS Lett. 311, 203-205, 1992

A/Title: p13, an endogenous target protein for arginine-specific ADP-ribosyltran

```

A:Reference number: S29229; MUID:93011981; PMID:1397316
A:Accession: S29229
A:Molecule type: protein
A:Residues: 24-46;69-100;116-130;143-166;200-213;219-319 <YAM>
C:Comment: This protein is a secretable component of promyelocyte granules. It is also c
C:Genetics:
A:Gene: mim-1
A>Note: this gene is activated by the product of the v-myb gene
C:Keywords: duplication
F:24-326/Product: myb-induced myeloid protein 1 #status experimental <MAT>

Query Match          53.7%; Score 433; DB 2; Length 326;
Best Local Similarity 50.3%; Pred. No. 4.7e-36;
Matches 80; Conservative 27; Mismatches 38; Indels 14; Gaps 2;

Oy 7 LLAGIISTALACD-----WANTCAGKSSNEITRTCDRHGCGYSQRSOR-P 52
      :|::||| |
Db 7 IALLSVSTFAFAROMEVHPPOQGRHMAQICSGNPFNRIRGDRYCYNVASRGGKEK 66
      ||::||| |
Oy 53 HGVVDILCSGSRYVAFTGTIVGOEKPYONKNAINNGVRISRCVMFEYTKPKYKG 112
      ||::||| |
Db 67 HGVVDICTDSGYVAFPSGGLSPFRFHNGNAIDDGQSISGSCYVKLVCIPHYRG 126
      ||::||| |
Oy 113 PIKGEKLGTLPLOKYVPYGIOSHVIENCSDSPTAVYL 151
      ||::||| |
Db 127 QIQKGQQQLGMRLPMQKVFPGLVSHIHVENCDQSDPTHL 165
      ||::||| |

RESULT 3
T1567
hypothetical protein K05F1.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T1567
R:Wohlmann, P.
submitted to the EMBL Data Library, June 1995
A:description: The sequence of C. elegans cosmid K05F1.
A:Reference number: T18537
A:Accession: T1567
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-305 <MOCH>
A:Cross-references: EMBL:029377; NID:g868173; PID:g868182; PIDN:AAA68719.1; CESP:K05F1.4
C:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:K05F1.4
A:introns: 44/2; 100/1; 164/1; 188/2; 295/1

Query Match          12.8%; Score 103.5; DB 2; Length 305;
Best Local Similarity 25.4%; Pred. No. 0.0087;
Matches 29; Conservative 22; Mismatches 36; Indels 27; Gaps 5;

Oy 55 GVDILCSAGSTVAPFTMGIMVGEEKPY--QNKNAINNGVRSIG----RGFCYKMFIKP 107
      |||::|::| |
Db 52 GDVVRCHLGEPRIYAPIGEMFYW-RPYGGRKREKSCADQGVRIEGTGOMOGYAVHISSVKL 110
      |||::|::| |
Oy 108 IKYKGPDIKKGEKLTPL---OKVYPGIOSHV-----HIENG 142
      ::|::|::| |
Db 111 SPFGGHVEADEIGEALNRYCFNDRGQNDVEPHVEIRLYKEERLDIPTHTLQNC 164

RESULT 4
C82840
conserved hypothetical protein XF0167 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 03-Jun-2002
C:Accession: C82840
R:anonymus, T.
Nature 406, 151-157, 2000
A:title: The genome sequence of the plant pathogen xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below

```

A:Accession: CB2840
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1,417 <SIN>
A:Cross-references: GB:AE003870; GB:AE003849; NID:g9104955; PIDN:AAF82980.1; GSP:
R:Experimental source: strain 945C
R:Simpsom, A.J.G.; Relnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarez
B:Rioses, M.R.S.; Bueno, M.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; C
as-Melo, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
Submitted to GenBank June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Ma
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmit
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.;
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da
M.; Tsunako, M.H.; Valada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore,
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
C:Genetics:
A:Gene: XP0167
C:Superfamily: hypothetical protein H10409
Query Match 10.0%; Score 80.5; DB 2; Length 417;
Best Local Similarity 25.0%; Pred. No. 2.5;
Matches 29; Conservative 24; Mismatches 42; Indels 21; Gaps 7;
OY 34 RCTDHGGCGYAAQSORPHOGVDILCSAGSYTAAFTGMI--VGQEKPRONKNAINGV 91
Db 249 RLSTFFGRRLHILGTSMRHKGVDAATAGTPTMAAGDARVQGTGCGKGVNYILHH-- 306
OY 92 RISGRFCVMEYIRPIKYKGRPKKGEKL--GTLPLQKYPYQIS----HWATE 140
Db 307 ---GKGY--RLTGKHSRF--GRIKAGQKINOGTVIG---YVGMTGLATGRPHLYAE 352
RESULT 5
F90571
Lipoprotein (Imported) - Mycoplasma pulmonis (strain UAB CT1P)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #extl-change 03-Aug-2001
R:Accession: F90571
R:Chambud, I.; Helliö, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Mos
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplas
A:Reference number: A9512; MUID:21267165; PMID:11353084
A:Accession: F90571
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-750 <KUN>
A:Cross-references: GB:AL445566; PID:g14089892; PIDN:CAJC3651.1; GSPDB:GN00153
A:Experimental source: strain UAB CT1P
C:Genetics:
A:Gene: MYPu_4780
A:Genetic code: SGC3
Query Match 9.8%; Score 79; DB 2; Length 750;
Best Local Similarity 22.8%; Pred. No. 6.4;
Matches 42; Conservative 25; Mismatches 55; Indels 62; Gaps 8;
OY 12 LISTAGPMAATCAGKSSNEIRTCDR--HGCQYAAQSORPHOGVDILCSAGSTYAP 69
Db 456 LQSPVSGWLD-----DRLSTYTSKADSKNTKDIHLGEDLILEQNKKEYIAP 502
OY 70 FTGMITVG--EKPYQ-----MKNAINNGVRIISGRFCVKMF-- 103
Db 503 FDGKIIASYYAPQAYGLGIVTVLEVMKKDLVQIDQIDQIAETDRIYIAFMNLN 562
OY 104 -----YIKPIKYKG-----PIKKEKIGTLLPLQKYPVGIQSHVIE--NCD 143
Db 563 PSYLENGYAEVQSSTAIEITPATPKTVKKEGVIG-LVGEKRNNGGMKPHVHIEVSLG 621

QY 144 SSDP 147
 DB 622 STNP 625

RESULT 6

hypothetical protein H10409 - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
 C:Accession: F64151
 R:Flerschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Guehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: F64151
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-475 <TRIG>
 A:Cross-references: GB:U32724; GB:U42023; NID:g1573378; PIDN:AA022068.1; PID:g1573382; T
 C:Superfamily: hypothetical protein H10409

Query Match

Best Local Similarity 9.6%; Score 77.5; DB 2; Length 475;
 Matches 27; Conservative 15; Mismatches 41; Indels 11; Gaps 4;

QY 51 RPHGVILCSAGTVVAPFTGMIVGOEKPYONKNAINGVRISGRGFCVAKMYIKP--I 108
 DB 346 RPHKGVDSVSQGPVIAPADGTV--EKVAYQAGAGRYVLMRGRGVYVYMLSKSLV 403
 QY 109 KYKGPVKKEKLTGLTLPLQKVPYPIQS--HVHTE 140
 DB 404 KAGQTVKKKERIAL-----SGNTGISTGPHLHTE 432

RESULT 7

hypothetical protein DKFZp566N1047.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
 C:Accession: T42644
 R:Ottewaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, November 1999
 A:Reference number: Z22231
 A:Accession: T42644
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-651 <AAB>
 A:Cross-references: EMBL:AL133109
 A:Experimental source: fetal kidney; clone DKFZp566N1047
 C:Genetics:
 A:Note: DKFZp566N1047.1

Query Match

Best Local Similarity 9.4%; Score 76; DB 2; Length 651;
 Matches 27; Conservative 14; Mismatches 25; Indels 20; Gaps 5;

QY 43 QYSAORSORPHGVILCSAGTVVAP-----FTGMIVGOEKPYONKNAINN--GVRI 93
 DB 500 QYNAORSQOMPQ-----AAQAGYQPVLSGGGQFQGLI-GVQCPPOSQNVINNOGGRPV 552
 QY 94 SGRGFCVAKMYIKPIKYKPIKKEK 119
 DB 553 QS-----VWVSYPMTSSYQVPMTOGSQ 574

RESULT 8

H84197
 hypothetical protein Vng0391c [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: H84197

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.;
 ; Jeltchauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe,
 A:Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: AB4160; MUID:20504483; PMID:11016950
 A:Accession: H84197
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-240 <STO>
 A:Cross-references: GB:AE004437; NID:g10580006; PIDN:AA018948.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: VNG0391C
 C:Superfamily: hypothetical protein MJ0570

Query Match
 Best Local Similarity 9.3%; Score 75; DB 2; Length 240;
 Matches 37; Conservative 23; Mismatches 56; Indels 42; Gaps 8;

QY 9 LAGLISTALAGPWANICAGKSSNEIRP-----CDRHGGGYSAORSORPHGVILCSA 62
 DB 98 VADTLDLDELGAG-----VIAGAVESDFQTSRMGMDRLGCELYAPLMQPPRELADAMTLA 153
 QY 63 GSTVYAPPTGCMV-----GQEKPYONKN-----ATNN--GVRISGRGFCVAKMYI 105
 DB 154 G-----FEIRITQYAAVYLDKSWLGRRLDRDLAALAAALDEYGVHMLGSGEFTLV 207
 QY 106 KPIKYPKIKKKEKLTGLTLPLQKVPYPIQS--HVHTE 142
 DB 208 DAPHMDRPIR-----LDYETVMAGDRHIETETAC 237

RESULT 9

hypothetical protein F9D16.290 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 17-Mar-2000
 C:Accession: T05614
 R:Bevan, M.; Weidner, H.; Weidner, E.; Wambutt, R.; Hoheisel, J.; Mewes, H.W.; May
 submitted to the Protein Sequence Database, February 1999
 A:Reference number: Z15419
 A:Accession: T05614
 A:Molecule type: DNA
 A:Residues: 1-444 <BEV>
 A:Cross-references: EMBL:AL035394
 A:Experimental source: cultivar Columbia; BAC clone F9D16
 C:Genetics:
 A:Map position: 4
 A:Insertions: 110/3; 151/1; 209/2; 233/1
 A:Note: F9D16.290
 C:Superfamily: polygalacturonase

Query Match

Best Local Similarity 9.3%; Score 75; DB 2; Length 444;
 Matches 39; Conservative 14; Mismatches 53; Indels 42; Gaps 8;

QY 23 NICAGKSSNEIRTC-----RHGGGYSAORSORPHGVILCSAGTVVAPFTGMIVGO 77
 DB 237 NVCI--EDSYISTGDDLVAKSGMDQYIAYG-RPSSNITIRRTGS---SPFAGIALGS 290
 QY 78 EKPYNKNAINNNGVRISGRGFCVAKMYI-----FYIKYK-----YKPIKKEKLG----- 121
 DB 291 ETSGIGKNIITAEHITLSNMGVGVNKTINIGRGVGIKIKIKISDYVDTAKYGIKIAGDTGD 350
 QY 122 -----TLPLQKVPYPIQS--HVHTE 141
 DB 351 HPDENYNPALPVR-----GIRIKN 371

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:10:39 ; Search time 8.5 Seconds

(without alignments)
736.815 Million cell updates/sec

Title: US-09-868-953a-2
Perfect score: 807
Sequence: 1 MFSTKALLAGLISALAGP.....GIQSHVHENCSSDPAYL 151

Scoring table: BLOSUM62
Gapop 10.0 , Gapect 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|---------------------|
| 1 | 806 | 99.9 | 151 | LCR2_HUMAN | 014960 homo sapien |
| 2 | 702 | 87.0 | 151 | LCR2_BOVIN | 062644 bos taurus |
| 3 | 685 | 84.9 | 151 | LCR2_MOUSE | 088803 mus musculu |
| 4 | 433 | 53.7 | 326 | MTM1_CHICK | P08940 gallus gall |
| 5 | 77.5 | 9.6 | 475 | YEBP_HAIEIN | P46633 haemophilus |
| 6 | 73 | 9.0 | 710 | DAS_PICAN | P06834 picchia anqu |
| 7 | 73 | 9.0 | 1330 | XDH_RAT | P22985 rattus norv |
| 8 | 72.5 | 9.0 | 355 | COL1_ARATH | 050055 arabidopsis |
| 9 | 70.5 | 8.7 | 625 | TRIL_MOUSE | 035305 mus musculu |
| 10 | 70.5 | 8.7 | 680 | Z334_HUMAN | 09hcz1 homo sapien |
| 11 | 69.5 | 8.6 | 472 | ZIM3_HUMAN | 09p6p6 homo sapien |
| 12 | 69 | 8.6 | 779 | TRF3_THEVO | 097a36 thermoplas |
| 13 | 68 | 8.4 | 184 | YTF63_CAERL | 011079 caenobacti |
| 14 | 68 | 8.4 | 762 | MUS2_AQUAE | 067287 aquifex aeo |
| 15 | 67.5 | 8.4 | 444 | PA7_RABIT | P98139 oryctolagus |
| 16 | 67.5 | 8.4 | 535 | VS50_HSVSA | 001012 herpesvirus |
| 17 | 66.5 | 8.2 | 1012 | DPOL_HSVJ | P52342 human herpe |
| 18 | 66 | 8.2 | 398 | MBL1_XENLA | P38565 xenopus lae |
| 19 | 66 | 8.2 | 564 | HEMA_IADA2 | P03446 influenza a |
| 20 | 65.5 | 8.1 | 362 | FIHP_ADE08 | P36845 human adeno |
| 21 | 65.5 | 8.1 | 466 | FA7_HUMAN | P08709 homo sapien |
| 22 | 65.5 | 8.1 | 1241 | NEPN_HUMAN | 060500 homo sapien |
| 23 | 65.5 | 8.1 | 1388 | RP0D_TOBAC | 029450 nicotiana t |
| 24 | 65 | 8.1 | 317 | IBP2_SHEEP | 099s65 staphylococ |
| 25 | 65 | 8.1 | 412 | GLYA_STAMP | 034498 clostridium |
| 26 | 64.5 | 8.0 | 357 | LEU3_CLOPA | P27113 oryctolagus |
| 27 | 64.5 | 8.0 | 391 | LEU2_BACSU | P27113 oryctolagus |
| 28 | 64.5 | 8.0 | 551 | LEU2_RABIT | P27113 oryctolagus |
| 29 | 64.5 | 8.0 | 2241 | TEGUF_HCVNA | P16785 human cytom |
| 30 | 64 | 7.9 | 271 | YLBG_ECOLI | P77518 escherichia |
| 31 | 64 | 7.9 | 466 | SELA_RHIME | P58262 rhizobium m |
| 32 | 64 | 7.9 | 593 | MTF2_HUMAN | O9Y483 homo sapien |
| 33 | 64 | 7.9 | 746 | GYP7_YEAST | P48365 saccharomyc |

ALIGNMENTS

| RESULT 1 | ID | LCR2_HUMAN | STANDARD: | PRT: | 151 AA. |
|----------|--|-----------------------------------|-----------|------|---------|
| AC | 014960 | 014565 | | | |
| DT | 15-JUL-1999 | (Rel. 38, Created) | | | |
| DT | 15-JUL-1999 | (Rel. 38, Last sequence update) | | | |
| DE | 16-OCT-2001 | (Rel. 40, Last annotation update) | | | |
| GN | Leukocyte cell-derived chemotaxin 2 precursor (hLECT2). | | | | |
| OS | Homo sapiens (Human) | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | | |
| OX | NCBI_TaxID:9606; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A., AND VARIANT ILE-58. | | | | |
| RC | TISSUE=Liver; | | | | |
| RX | MEDLINE=98193133; PubMed=9524238; | | | | |
| RA | Yamagoe S., Mizuno S., Suzuki K.; | | | | |
| RT | "Molecular cloning of human and bovine LECT2 having a neutrophil | | | | |
| RT | chemotactic activity and its specific expression in the liver."; | | | | |
| RL | Biochim. Biophys. Acta 1396:105-113(1998). | | | | |
| RN | [2] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | TISSUE=Peripheral blood; | | | | |
| RX | MEDLINE=98207247; PubMed=9545637; | | | | |
| RA | Yamagoe S., Kameoka Y., Hashimoto K., Mizuno S., Suzuki K.; | | | | |
| RT | "Molecular cloning, structural characterization, and chromosomal | | | | |
| RT | mapping of the human LECT2 gene."; | | | | |
| RL | Genomics 48:324-329(1998). | | | | |
| RN | [3] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RX | MEDLINE=99048897; PubMed=9832057; | | | | |
| RA | Nagai H., Hamada T., Uchida T., Yamagoe S., Suzuki K.; | | | | |
| RT | "Systemic expression of a newly recognized protein, LECT2, in the | | | | |
| RT | human body."; | | | | |
| RL | Pathol. Int. 48:882-886(1998). | | | | |
| RN | [4] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RA | Kimmerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M., | | | | |
| RA | Kadner K., Miguel T., Miller C., Pittluck S., Pollard M., Rojeski H., | | | | |
| RA | Subramanian S., Martin C.H.; | | | | |
| RL | Submitted (May-1998) to the EMBL/GenBank/DBJ databases. | | | | |
| RN | [5] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RA | Strong C., Biewald T., Tin-Wollam A., Duckels G.; | | | | |
| RA | Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases. | | | | |
| RN | [6] | | | | |
| RP | SUBCELLULAR LOCATION. | | | | |
| RX | MEDLINE=97031488; PubMed=8877413; | | | | |
| RA | Yamagoe S., Yamakawa Y., Matsuo Y., Minowada J., Mizuno S., Suzuki K.; | | | | |
| RT | "Purification and primary amino acid sequence of a novel neutrophil | | | | |
| RT | chemotactic factor LECT2."; | | | | |
| RL | Immunol. Lett. 52:9-13(1996). | | | | |
| CC | -!- FUNCTION: Has a neutrophil chemotactic activity. Also a positive | | | | |
| CC | regulator of chondrocyte proliferation. Also a positive | | | | |
| CC | -!- SUBCELLULAR LOCATION: Cytoplasmic and secreted. | | | | |

CC -1- TISSUE SPECIFICITY: Highly expressed in adult and fetal liver and
 CC weakly in testis. Not expressed in bone marrow.
 CC -1- INDUCTION: By phytohemagglutinin (PHA).
 CC -1- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D63521; BAA23609.1; -
 DR EMBL: AB007546; BAA25669.1; -
 DR EMBL: AC004763; AAC17734.1; -
 DR EMBL: AC002428; AAB66905.1; -
 DR Genew; HGNC:6550; LECT2.
 DR MIM; 602882; -
 KW Chemotaxis; Signal; Polymorphism.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT CHAIN 19 151 LEUKOCYTE CELL-DERIVED CHEMOTAXIN 2.
 FT VARIANT 58 58 V->I
 FT SEQUENCE 151 AA; 16376 MW; A4267F0A24E69631 CRC64;
 SQ
 Query Match 99.9%; Score 806; DB 1; Length 151;
 Best Local Similarity 99.3%; Pred. No. 2.4e-76;
 Matches 150; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MESTKALLAGLISTALAGPWANICAGKSSNEIRTCDRHGGCGOYSQRSORPHQGVYDLC 60
 DB 1 MESTKALLAGLISTALAGPWANICAGKSSNEIRTCDRHGGCGOYSQRSORPHQGVYDLC 60
 QY 61 SAGSTYVAPFTGMIVGGEKPYQNNKAINNGVRIISGRCVYKMFYIKPIYKGIKKGKGL 120
 DB 61 SAGSTYVAPFTGMIVGGEKPYQNNKAINNGVRIISGRCVYKMFYIKPIYKGIKKGKGL 120
 QY 121 GTLLPLQKVPYGIQSHVHIENCDSDPTAYL 151
 DB 121 GTLLPLQKVPYGIQSHVHIENCDSDPTAYL 151
 RESULT 2
 LCT2_BOVIN STANDARD; PRT; 151 AA.
 ID LCT2_BOVIN
 AC 062644;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leukocyte cell-derived chemotaxin 2 precursor (bLECT2) (Chondromodulin
 DE II) (bcHM-II).
 GN LECT2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=98193133; PubMed=9524238;
 RA Yamagoe S., Mizuno S., Suzuki K.;
 RT "Molecular cloning of human and bovine LECT2 having a neutrophil
 RT chemotactic activity and its specific expression in the liver.";
 RL Biochim. Biophys. Acta 1396:105-113(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=99160594; PubMed=10050029;
 RA Shukunami C., Kondo J., Wakai H., Takahashi K., Inoue H., Kamizono A.,
 RA Hiraki Y.;
 RT "Molecular cloning of mouse and bovine chondromodulin-II cDNAs and the

RT growth-promoting actions of bovine recombinant protein.";
 RL J. Biochem. 125:436-442(1999).
 RN [3]
 RP SEQUENCE OF 19-151.
 RC TISSUE=Epiphyseal cartilage;
 RX MEDLINE=96394331; PubMed=8796437;
 RA Hiraki Y., Inoue H., Kondo J., Kamizono A., Yoshitake Y.,
 RA Shukunami C., Suzuki F.;
 RT "A novel growth-promoting factor derived from fetal bovine cartilage,
 RT chondromodulin II. Purification and amino acid sequence.";
 RL J. Biol. Chem. 271:22657-22662(1996).
 CC -1- FUNCTION: Has a neutrophil chemotactic activity. Also a positive
 CC regulator of chondrocyte proliferation.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB001350; BAA25302.1; -
 DR EMBL: D89011; BAB18616.1; -
 KW Chemotaxis; Signal.
 FT SIGNAL 1 18 LEUKOCYTE CELL-DERIVED CHEMOTAXIN 2.
 FT CHAIN 19 151
 FT SEQUENCE 151 AA; 16320 MW; 2A1DB8F5B28A3D8 CRC64;
 SQ
 Query Match 87.0%; Score 702; DB 1; Length 151;
 Best Local Similarity 86.1%; Pred. No. 1.4e-65;
 Matches 130; Conservative 8; Mismatches 13; Indels 0; Gaps 0;
 QY 1 MESTKALLAGLISTALAGPWANICAGKSSNEIRTCDRHGGCGOYSQRSORPHQGVYDLC 60
 DB 1 MESTGTLALALISPALAGPAITCAGKSSNEIRTCDRHGGCGOYTRQRNKLHQGVYDLC 60
 QY 61 SAGSTYVAPFTGMIVGGEKPYQNNKAINNGVRIISGRCVYKMFYIKPIYKGIKKGKGL 120
 DB 61 SAGSTYVAPFTGMIVGGEKPYKNNKAINNGVRIISGRCVYKMFYIKPIYKGIKKGKGL 120
 QY 121 GTLLPLQKVPYGIQSHVHIENCDSDPTAYL 151
 DB 121 GTLLPLQKVPYGIQSHVHIENCDSDPTAYL 151
 RESULT 3
 LCT2_MOUSE STANDARD; PRT; 151 AA.
 ID LCT2_MOUSE
 AC 088803; O88804; O9QWN3; O9Z337;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leukocyte cell-derived chemotaxin 2 precursor (Chondromodulin II)
 DE (ChM-II).
 GN LECT2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANT VAL-129.
 RC STRAIN=BALB/c; TISSUE=Liver;
 RX MEDLINE=98382586; PubMed=9714793;
 RA Yamagoe S., Watanabe T., Mizuno S., Suzuki K.;
 RT "The mouse lect2 gene: cloning of cDNA and genomic DNA, structural
 RT characterization and chromosomal localization.";
 RL Gene 216:171-178(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=Swiss Webster / NIH Swiss; TISSUE=Embryo, and Liver;


```

RA MEDLINE-99160594; PubMed-10050029;
RX Shukunami C., Kondo J., Wakai H., Takahashi K., Inoue H., Kamizono A.,
RA Hiraki Y.;
RT "Molecular cloning of mouse and bovine chondromodulin-II cDNAs and the
RT growth-promoting actions of bovine recombinant protein.";
RL J. Biochem. 123:436-442(1999).
CC -I- FUNCTION: Has a neutrophil chemotactic activity. Also a positive
CC regulator of chondrocyte proliferation.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC CC -I- ALTERNATIVE PRODUCTS: 2 isoforms, 1/LECT2 (shown here) and
CC 2/LECT20; are produced by alternative splicing.
CC -I- TISSUE SPECIFICITY: Highly expressed in liver and weakly in
CC testis. Not expressed in heart, brain, spleen, lung, skeletal
CC muscle and kidney.
CC CC -I- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch)
-----
DR DR EMBL: AB009687; BAA33383.1; -
DR DR EMBL: AB009688; BAA33384.1; -
DR DR EMBL: AB009689; BAA33385.1; -
DR DR EMBL: AB009689; BAA33386.1; -
DR DR EMBL: AF035161; AAP13302.1; -
DR DR MGI: 1278342; Lect2.
KW Chemotaxis; Signal; Alternative splicing.
FT SIGNAL 1 18
FT CHAIN 19 151 LEUCOCYTE CELL-DERIVED CHEMOTAXIN 2.
FT FT FCVDFIFKPIKKGSIKSGEIKGTLLPDKRGIGSHV
FT VARSPLIC 98 151 VENCDSDPTALV->OQLAHTLVNVTVCWDKIQIPR
FT PTRELECOMFLH (IN ISOFORM LECT20).
FT FT
FT VARIANT 129 129 I->V
SQ SEQUENCE 151 AA; 16405 MW; 1BAFA44046B7AE8E CRC64;
Query Match 84.9%; Score 685; DB 1; Length 151;
Best Local Similarity 82.1%; Pred. No. 7.8e-64;
Matches 124; Conservative 13; Mismatches 14; Indels 0; Gaps 0;
QY 1 MESTKALLAGLSTALAGPWANICAGKSNEIRTCRDRCGCGYSNRQRPHGVILC 60
DB | | : | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 MIPTTLISALSSALAGPWANICASSKSNEIRTCDSYGCGYSAQRTGRHNHGCVYL 60
QY 61 SAGSYVAPETGMIVGEKPYNKNNAINGVAISRGFCVMKYIRPKKKGPTRKEKL 120
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 61 SDGSVVYAPETGKIVGEGEKPRNNKNAINDGIRLSGFGCFVKIRPKRKGSITKKEKL 120
QY 121 GTLLPLQKVYPGIGIOSHYHIENCDDSDPYAL 151
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 121 GTLLPLQKVYPGIGIOSHYHIENCDDSDPYAL 151
RESULT 4
MIM1_CHICK
ID ID MIM1_CHICK STANDARD; PRT; 326 AA.
AC AC P08940;
DT DT 01-AUG-1990 (Rel. 15, Created)
DT DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE DE Myeloid protein-1 precursor (P33).
GN GN MIM-1.
OS OS Gallus gallus (Chicken).
OC OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
OC Archosauroidae; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX OX NCBI_TaxID=9031;
RN RN [1]
RP RP SEQUENCE FROM N.A.
RX RX MEDLINE-90090611; PubMed-2688896;
```

```

RA Ness S.A.; Marknell A., Graf T.;
RA "The v-myb oncogene product binds to and activates the promyelocyte-
RT specific m1m-1 gene".
RL Cell 59:1115-1125(1989).
[2]
RN PARTIAL SEQUENCE, AND SEQUENCE OF 295-299 FROM N.A.
RP STRAIN-White leghorn;
RC MEDLINE=93011981; PubMed=1397316;
RX Yamada K., Tsuchiya M., Mishima K., Shimoyama M.;
RA "p23, an endogenous target protein for arginine-specific ADP-
RT ribosyltransferase in chicken polymorphonuclear leukocytes, is highly
RT homologous to m1m-1 protein (myb-induced myeloid protein-1).";
RL FEBS Lett. 311:203-205(1992).
CC -1- FUNCTION: NOT KNOWN.
CC -1- SUBCELLULAR LOCATION: GRANULES OF PROMYELOCYTES.
CC -1- INDUCTION: BY THE MYB ONCOGENE.
CC -1- PFM: SUBSTRATE FOR ARGININE-SPECIFIC ADP-RIBOSYLTRANSFERASE.
CC -1- SIMILARITY: BELONGS TO THE LECT2 / M1M-1 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.sib.ch).
-----
DR EMBL; M29448; AAA48954.1; -
DR EMBL; M29449; AAA48958.1; -
DR PIR; A33755; A33755.
DR PIR; S29229; S29229.
KM Repeat; Signal.
FT SIGNAL 1 23 PROBABLE.
FT CHAIN 24 326 MYELOID PROTEIN-1.
FT REPEAT 28 162 1.
FT REPEAT 177 312 2.
FT CONFLICT 297 297 1 -> Y (IN REF. 1)
SQ SEQUENCE 326 AA; 35636 MW; B8C4742EF0D3BAD3 CRC64;
Query Match 53.7%; Score 433; DB 1; Length 326;
Best Local Similarity 50.3%; Pred. No. 2e-37;
Matches 80; Conservative 27; Mismatches 38; Indels 14; Gaps 2;
QY 7 LLAGLISYALAP-----NANICAGKSNERTCDRHCCGGVSAQR--P 52
Db 7 IALLSVSTAFRAQOMEVHPQQGRHWQIOTSGNPNRIRGCDRYCGNGASQGGK 66
QY 53 HGGVDILCSGSTVYAFPTGMIYGOEKPYQNNKAINNGVRIISGRGCVKMFYIKPIRYK 112
Db 67 HKGVDVICTDGSIVYAFPSQGLSGPIREFHNGNALIDGVQISGSGVCVKRWCHPIRYHG 126
QY 113 PIKKGKLGTLPLQKYRPGIQSHVHIFNCSSDPTAYL 151
Db 127 OTQGGQGLSRMLPMQKVPGLIVSHIHVENCDSDDPTHL 165
RESULT 5
YERB_HAEIN
ID YERB_HAEIN STANDARD; PRT; 475 AA.
AC P44693;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein HI0409.
GN HI0409.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;

```

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Greth C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.,
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: COULD BE INVOLVED IN CELL WALL DEGRADATION OR
CC FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: STRONG TO E. COLI YEBA
CC -1- SIMILARITY: TO STAPHYLOCOCCUS LYSOSTAPHIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U32724; AAC22068.1; .
CC DR EMERP: M37 UPB; .
CC DR TIGR: H10409; .
CC DR InterPro: IPR002886; Peptidase_M37.
CC DR Pfam: PF01551; Peptidase_M37; 1.
CC DR Hypothetical protein; Cell wall; Complete proteome.
CC KW SEQUENCE 475 AA; 53255 MW; 143C10F92233939D CRC64;
SQ
Query Match 9.6%; Score 77.5; DB 1; Length 475;
Best Local Similarity 28.7%; Pred. No. 1.7; Indels 11; Gaps 4;
Matches 27; Conservative 15; Mismatches 41;
QY 51 RPHGVLDILCSAGTYAFETGMIVGOEKPYONKNAINNGVRISGRFCVKNFYIKP--I 108
DB 346 RPHKVDVSVSGTPIYADGTIV--EKVAYVAGAGRYVALRGRGREGYQYVMHLSKSLV 403
QY 109 KYKGPVKGEKLGTLPLQKVPYGIOS--HVHIE 140
DB 404 KAGGTAKKGERIAL-----SGNTGISGPHLHYE 432
RESULT 6
DAS_PICAN STANDARD; PRT; 710 AA.
AC P06834;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Dihydroxy-acetone synthase (EC 2.2.1.3) (DHAS) (formaldehyde
DE transketolase) (glycerone synthase).
GN DAS.
OS Pichia angusta (Yeast) (Hansenula polymorpha).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4905;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 34438;
RX MEDLINE=85215670; PubMed=2987872;
RA Janowicz Z.A., Eckart M.R., Drewek C., Roggenkamp R.O.,
RA Hollenberg C.P., Maat J., Ledebor A.M., Visser C., Verrips C.T.;
RT "Cloning and characterization of the DAS gene encoding the major
RT methanol assimilatory enzyme from the methylotrophic yeast Hansenula
RT polymorpha.";
RL Nucleic Acids Res. 13:3043-3062(1985).
RN [2]
RP REVISIONS: SEQUENCE OF 667-710 FROM N.A.
RC STRAIN=ATCC 34438;

RX MEDLINE=93101130; PubMed=1465101;
RA Hansen H., Didion T., Thiemann A., Veenhuis M., Roggenkamp R.O.;
RT "targeting sequences of the two major peroxisomal proteins in the
RT methylotrophic yeast Hansenula polymorpha.";
RL Mol. Gen. Genet. 235:269-278(1992).
CC -1- FUNCTION: THIS IS THE MAJOR METHANOL ASSIMILATORY ENZYME FROM
CC THE METHYLOTROPHIC HANSENULA POLYMORPHA.
CC CATALYTIC ACTIVITY: D-xylulose 5-phosphate + glyceraldehyde -
CC glyceraldehyde 3-phosphate + glyceraldehyde -
CC COFACTOR: THIAMINE PYROPHOSPHATE.
CC -1- SUBCELLULAR LOCATION: Peroxisomal.
CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X02424; CA26276.1; .
CC DR PIR: A23009; X1HOFK.
CC DR HSP: P23254; ITRK.
CC DR InterPro: IPR000360; Transketolase.
CC DR Pfam: PF00456; transketolase; 1.
CC DR Pfam: PF02779; transket-pyr; 1.
CC DR PROSITE: PS00801; TRANSKETOLASE_C; 1.
CC DR PROSITE: PS00802; TRANSKETOLASE_2; 1.
CC DR PROSITE: PS00342; MICRORODIES_CTER; 1.
CC DR Methanol utilization; Peroxisome; Transferrase; thiamine pyrophosphate.
CC KW SITE 708 710 MICRORODIES TARGETING SIGNAL (POTENTIAL).
CC FT SEQUENCE 710 AA; 78842 MW; E4424BD37CC8B0B CRC64;
SQ
Query Match 9.0%; Score 73; DB 1; Length 710;
Best Local Similarity 24.1%; Pred. No. 7.6; Indels 34; Gaps 7;
Matches 33; Conservative 21; Mismatches 49;
QY 11 GLISTALGPMANICAGKSSNEIRTCDRHSCGYSQAORSRPHGVLDILCSAGTYVAPF 70
DB 525 GLFOKAVELPFSISLS-LRNEVL-----QYLABRAQRNRNAGYILEDNAEAVQI 575
QY 71 TGMIVGOEKPYONKNAINNGVRISGRF-----CYKMYPIPIKIKGP--KKEGL 120
DB 576 IG--VGAEMEFADKKA-----KILGRFTRVLSTPCTRLFDEOSIGYRSVLRDGRV 628
QY 121 GTLLPLQKVPYGIOSHV 137
DB 629 PTVV-----VDGHV 637
RESULT 7
XDH_RAT STANDARD; PRT; 1330 AA.
AC P22985; Q63157;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Xanthine dehydrogenase/oxidase [includes: Xanthine dehydrogenase
DE (EC 1.1.1.204) (XD); Xanthine oxidase (EC 1.1.3.22) (XO) (Xanthine
DE oxidoreductase)].
GN XDH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=90354396; PubMed=2387845;
RA Amaya Y., Yamazaki K.-I., Sato M., Noda K., Nishino T., Nishino T.;
RT "Proteolytic conversion of xanthine dehydrogenase from the

RT NAD-dependent type to the O2-dependent type. Amino acid sequence of
 RT rat liver xanthine dehydrogenase and identification of the cleavage
 RT sites of the enzyme protein during irreversible conversion by
 RT trypsin.";
 RT J. Biol. Chem. 265:14170-14175(1990).
 RN [2]
 RP SEQUENCE OF 1-54 FROM N.A.
 RC STRAIN-Sprague-Dawley;
 RX MEDLINE=94268906; PubMed=8208609;
 RA Chow C.W., Clark M., Rinaldo J., Chalkley R.;
 RT "Identification of the rat xanthine dehydrogenase/oxidase promoter.";
 RT Nucleic Acids Res. 22:1846-1854(1994).
 CC -1- FUNCTION: THIS ENZYME CAN BE CONVERTED FROM THE DEHYDROGENASE FORM
 CC (D) TO THE OXIDASE FORM (O) IRREVERSIBLY BY PROTEOLYSIS OR
 CC REVERSIBLY THROUGH THE OXIDATION OF SULFHYDRYL GROUPS.
 CC -1- CATALYTIC ACTIVITY: Xanthine + NAD(+) + H(2)O = urate + NADH.
 CC -1- CATALYTIC ACTIVITY: Xanthine + H(2)O = urate + H(2)O(2).
 CC -1- COFACTOR: FAD, MOLIBDOPTERIN, AND TWO 2FE-2S CLUSTERS.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Peroxisomal.
 CC -1- INDUCTION: BY INTERFERON.
 CC -1- SIMILARITY: BELONGS TO THE XANTHINE DEHYDROGENASE FAMILY.
 CC -1- SIMILARITY: TO 2FE-2S FERRODOXINS IN THE N-TERMINAL DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J05579; AAA42349.1; .
 DR EMBL: U08122; AAA18869.1; .
 DR EMBL: U08120; AAA18869.1; JOINED.
 DR EMBL: U08121; AAA18869.1; JOINED.
 DR PIR: A37810; A37810.
 DR HSSP: P80457; IPIQ.
 DR InterPro: IPR002888; 2Fe-2S bind.
 DR InterPro: IPR000564; 2Fe2S ferredoxin.
 DR InterPro: IPR000674; Aldxan_dh_C.
 DR InterPro: IPR005107; CO_dh_flav_C.
 DR InterPro: IPR000572; Euk_Mb_oxred.
 DR InterPro: IPR001041; Ferredoxin.
 DR InterPro: IPR002346; dehydrog_molypb.
 DR Pfam: PF00111; fer2; 1.
 DR Pfam: PF00941; FAD_binding_5; 1.
 DR Pfam: PF01315; Ald_xan_dh_C; 1.
 DR Pfam: PF01799; fer2_2; 1.
 DR Pfam: PF02738; Ald_xan_dh_C2; 1.
 DR Pfam: PF03450; CO_dh_flav_C; 1.
 DR Prodom: PD16071; 2Fe-2S bind; 1.
 DR PROSITE: PS00197; 2FE2S_FERREDOXIN; 1.
 DR PROSITE: PS00559; MOLIBDOPTERIN_EUK; 1.
 DR Oxidoreductase; NAD; Molybdenum; Flavo-protein; FAD; Iron-sulfur.
 FT INIT_MET 0
 FT METAL 36 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 42 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 47 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 50 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 SQ SEQUENCE 1330 AA; 146111 MW; A3DD206B9D74E565 CRC64;
 Query Match 9.0%; Score 73; DB 1; Length 1330;
 Best local Similarity 28.8%; Pred. No. 15;
 Matches 34; Conservative 12; Mismatches 42; Indels 30; Gaps 6;
 OY 35 TCDRHGCGQYSA--QSRPHQGVADILCSAGSTVYAFPTGMVQGEKPYONKNAINGV 91
 Db 660 TCVCIIIGAVVADPDEHQAARGVKI-----TYEDPALITID-----AINN-- 703
 OY 92 RISRGFVVMKFKPIKX-KGPIKGEKGLTLLPLQKVPYGIOASHWIE-NCDSDDP 147
 Db 704 -----NSFYGESEIKIEKGLDKKGFSEADNVVSGELYIGQEHFLEETNCTIAPV 752

RESULT 8
 COLL_ARATH
 ID COLL_ARATH STANDARD; PRT; 355 AA.
 AC 050035;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Zinc finger protein CONSTANS-LIKE 1.
 GN COL1 OR AR5G15850 OR F14F8_230.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Landsberg erecta; TISSUE-Aerial parts;
 RA Putterill J.J., Ledger S.E., Lee K., Robson F., Murphy G.,
 RA Coupland G.;
 RT "The flowering-time gene CONSTANS and homologue CONSTANS LIKE 1 exist
 RT as a tandem repeat on chromosome 5 of Arabidopsis.";
 RT (in) Plant Gene Register PCR97-077.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tsubata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsunoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naito K., Okumura S., Shino S., Takeuchi G., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Huang E., Spiegel L., Gao J., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoneking T., Peplin K., Spleth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mlvaney E., Ozersky P., Riley A., Stromwallt C.,
 RA Wagner-McPherson C., Wollam A., Yoakum M., Bell M., Dedha N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchhoff K., Toth K., King L., Bahnet A., Miller B., Marra M.,
 RA Markissen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wambutt R., Duesterhoeft A., Stekema W., Pohl T.,
 RA Ertlan K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Gymnopoulos B., Zimmermann W.,
 RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirkse W., Mooljman P., Klein Lankhorst R.,
 RA Weltzenegger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villarroel R., Giesen J., Ardiles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K., Rudd S., Schoof H.,
 RA Scheller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:823-826(2000).
 RN [3]
 RP CHARACTERIZATION.
 RC STRAIN-cv. C24;
 RX MEDLINE=21260187; PubMed=11359606;
 RA Ledger S., Strayer C., Ashton F., Kay S.A., Putterill J.,
 RT "Analysis of the function of two circadian-regulated CONSTANS-LIKE
 RT genes.";
 RT Plant J. 26:15-22(2001).
 CC -1- FUNCTION: Putative transcription factor that may be involved in
 CC the light input to the circadian clock but does not affect
 CC flowering time.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LEAVES AND AT LOWER LEVELS
 CC IN STEMS, FLOWERS AND SILIQUES, NOT DETECTED IN ROOTS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT.
 CC -1- INDUCTION: EXPRESSED WITH A CIRCADIAN RHYTHM SHOWING A PEAK
 CC AT DAWN.
 CC -1- SIMILARITY: BELONGS TO THE CONSTANS FAMILY.

```

CC -1- SIMILARITY: CONTAINS 2 B BOX-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y10555; CAA71587.1; -
DR EMBL: Y10556; CAA71588.1; -
DR EMBL: AL391144; CAC01784.1; -
DR InterPro: IPR000315; znf_Box.
DR InterPro: IPR002926; znf_constants.
DR Pfam: PF00643; zf-B_box; 2.
DR ProDom: PD007661; znf_constants; 2.
DR SMART: SM00336; BBOX; 1.
DR PROSITE: PS01119; zf_BBOX; 2.
KW Transcription regulation; Zinc-finger; DNA-binding; Nuclear protein;
KW Repeat; Multigene family.
FT ZN_FING 12 54 B BOX-TYPE 1.
FT ZN_FING 55 97 B BOX-TYPE 2.
SQ SEQUENCE 355 AA; 39491 MW; 412C237A9CD30309 CRC64;

Query Match
Best Local Similarity 24.4%; Score 72.5; DB 1; Length 355;
Matches 30; Conservative 15; Mismatches 43; Indels 35; Gaps 4;

QY 20 PANNICAGKSSNEIRTCDRHCGGQYSAGRSQRPHGVDL-----CSAGSTV 66
DB 62 PAAPFCADASLCTTCDSE---PYONKNAINGVIRISGRCFCVKMEYIKPIKYGK 118
QY 67 YAPFTGMIYGOEK-----PYONKNAINGVIRISGRCFCVKMEYIKPIKYGK 113
DB 119 TDENRLVVLQGEDEDEAEAAASMLPNSGKNSGNN-----NGFSIGDEFLNLYDSSS 172
QY 114 IKR 116
DB 173 DKQ 175

RESULT 9
TR11_MOUSE STANDARD; PRT; 625 AA.
ID TR11_MOUSE
AC O35305; Q8VCT7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily 11A precursor
DE (Receptor activator of NF-kB) (Osteoclast differentiation factor
DE receptor) (ODPR).
GN TNFRSF11A OR RANK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retal Liver;
RX MEDLINE=98032977; PubMed=9367155;
RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougal W.C.,
RA Tometsko M.E., Roux E.R., Teepe M.C., Dubose R.F., Cosman D.,
RA Galibert L.;
RA "A homologue of the TNF receptor and its ligand enhance T-cell growth
RA and dendritic-cell function.";
RT Nature 390:175-179(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strausberg R.;
RU Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [3]

```

```

RP FUNCTION.
RX MEDLINE=9097247; PubMed=9878548;
RA Nakagawa N., Kinoshita M., Yamaguchi K., Shima N., Yasuda H., Yano K.,
RA Morinaga T., Higashio K.;
RT "RANK is the essential signaling receptor for osteoclast
RT differentiation factor in osteoclastogenesis.";
RL Biochem. Biophys. Res. Commun. 253:395-400(1998).
CC -1- FUNCTION: Receptor for TNFRSF11/RANKL/TRANCE/OPGL; essential for
CC RANKL-mediated osteoclastogenesis. Involved in the regulation of
CC interactions between T-cells and dendritic cells.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: UBIQUITOUS EXPRESSION WITH HIGH LEVELS IN
CC TRABECULAR BONE, THYMUS, SMALL INTESTINE, LUNG, BRAIN AND KIDNEY.
CC WEAKLY EXPRESSED IN SPLEEN AND BONE MARROW.
CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF019046; AAB86810.1; -
DR EMBL: BC019185; AAH19185.1; -
DR HSSP: P25942; ICDF.
DR MGD: MGI:1314891; Tnfrsf11a.
DR InterPro: IPR001368; TNFR_c6.
DR Pfam: PF00020; TNFR_c6; 3.
DR ProDom: PD000771; TNFR_c6; 1.
DR SMART: SM00208; TNFR; 4.
DR PROSITE: PS00652; TNFR_NGFR_1; 1.
DR PROSITE: PS00500; TNFR_NGFR_2; 1.
KW Receptor; Transmembrane; Glycoprotein; Repeat; signal.
FT SIGNAL 1 30
FT CHAIN 31 625
FT FT 31 625
FT DOMAIN 31 214
FT TRANSMEM 215 234
FT DOMAIN 235 625
FT REPEAT 35 69
FT REPEAT 72 113
FT REPEAT 115 152
FT REPEAT 155 195
FT DISULFID 35 47
FT DISULFID 48 61
FT DISULFID 51 69
FT DISULFID 72 87
FT DISULFID 93 113
FT DISULFID 115 128
FT DISULFID 134 152
FT CARBOHYD 106 106
FT CARBOHYD 175 175
FT CONFLICT 494 494
SQ SEQUENCE 625 AA; 66621 MW; F8C1872B95511D8E CRC64;

Query Match
Best Local Similarity 8.7%; Score 70.5; DB 1; Length 625;
Matches 31; Conservative 12; Mismatches 39; Indels 63; Gaps 5;

QY 16 ALAGPMANIT-----CAGKSSNEIRTCRHHGCGQYSAGRSQRPHGVDILCSA 62
DB 309 AAGGPMAEVDRSRTFLVSEVETGDLRSKRIPTDEY-----TDRPSQPSFSLLIQ 362
QY 63 GSTRVAPFTGMIYGOEKPYONKNAINNVIRISGRCFCVKMEYIKPIKYGKELGT 122
DB 363 GSKSIPF-----QE-----PLEVGENDS- 381
QY 123 LPLQKVPYPGIQSHVHENCSDSDP 147
DB 382 ----LSQCFGTGTSTVDSGCDFTPEP 403

```


CC STRAIN-Bristol N2;
 RA Favello T.;
 RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE BIL FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U28740; AAA68321.1; -
 DR Wormpep: B0563.3; CE02441.
 DR InterPro: IPR002199; Bax_inhbrtl.
 DR Pfam: PF01027; UPF0005; 1.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 70 90 POTENTIAL.
 FT TRANSMEM 96 116 POTENTIAL.
 FT TRANSMEM 123 143 POTENTIAL.
 FT TRANSMEM 160 180 POTENTIAL.
 SQ SEQUENCE 184 AA; 20942 MW; 09836D218202F378 CRC64;
 Query Match 8.4%; Score 68; DB 1; Length 184;
 Best Local Similarity 27.2%; Pred. No. 5.9;
 Matches 31; Conservative 9; Mismatches 32; Indels 42; Gaps 7;
 QY 25 CAKSSNEIRTCDRHGGCGQYSAQ--RSQRPFGVDILCSAGS-----TVY--- 67
 Db 21 CVPKS-----RFGNSDYCAPRLRQSPSQ---LCSFGRTHRSSCHYGMULFEAK 69
 QY 68 -----APFTGMIVGOEKPY--QNKNAINNGVRSIGRGFCV-----KMFYIKP 107
 Db 70 VLEAAVITGLVVASLFRATYITQNKRPDSVGASMGSLCYLLMAGITQMFMPSP 123
 RESULT 14
 MUS2_AOUAE STANDARD; PRT; 762 AA.
 ID MUS2_AOUAE
 AC 067287;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Muts2 protein.
 GN MUTSB OR MUTS2 OR AQ_1242.
 OS Aquifex aeolicus.
 CC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
 CC Aquifex.
 CC NCBI_TaxID=63363;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Feldman D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus";
 RT Nature 392:353-358(1998).
 CC -1- FUNCTION: NOT KNOWN.
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE00730; AAC07247.1; -

DR InterPro: IPR000302; KOM_motif.
 DR InterPro: IPR000432; Muts_C.
 DR InterPro: IPR002863; Muts_N.
 DR InterPro: IPR002625; Smr/Muts2.
 DR Pfam: PF00467; KOM; 1.
 DR Pfam: PF00488; Muts_C; 1.
 DR Pfam: PF01713; Smr; 1.
 DR ProDom: PD001263; Muts_C; 1.
 DR SMART: SM00534; MutsSac; 1.
 DR SMART: SM00533; Mutsd; 1.
 DR SMART: SM00463; Smr; 1.
 DR TIGRfams: TIGR01069; muts2; 1.
 DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; FALSE_NEG.
 KW ATP-binding; DNA-binding; Complete proteome.
 FT NP_BIND 329 336 ATP (POTENTIAL).
 SQ SEQUENCE 762 AA; 87538 MW; 3A7D977DC2EB01BC CRC64;
 Query Match 8.4%; Score 68; DB 1; Length 762;
 Best Local Similarity 21.8%; Pred. No. 27;
 Matches 27; Conservative 19; Mismatches 40; Indels 38; Gaps 4;
 QY 64 STYVAPFTGMIVGOEKPYQNKNAINNGVRSIGRGF-----CYKMYIRPIKKGPIKGEK 119
 Db 20 SYVSPAPKTKETQNLKPYTNKKEVKEIELSAFFIAENVRLEFEEDIRL--LTKAKL 77
 QY 120 LGTL-----PLQKYPGIGSHVHLEN--CDSSDP 147
 Db 78 QGALGVEDILKILNVILTKIRRLVSLSSHVQRLEPLRKYYKKLYTSPLNLIGSIDP 137
 QY 148 TAYL 151
 Db 138 RGFV 141
 RESULT 15
 FA7_RABIT STANDARD; PRT; 444 AA.
 ID FA7_RABIT
 AC P98139; P79224;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Coagulation factor VII precursor (EC 3.4.21.21) (Serum prothrombin
 DE conversion accelerator).
 GN F7.
 OS Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 CC NCBI_TaxID=9986;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93190306; PubMed=833365;
 RA Brothers A.B., Clarke B.J., Sheffield W.P., Blajchman M.A.;
 RT "Complete nucleotide sequence of the cDNA encoding rabbit coagulation
 RT factor VII";
 RT Thromb. Res. Suppl. 69:231-238(1993).
 CC [2]
 CC REVISION TO 395.
 RC TISSUE=Liver;
 RA Ruiz S.R., Blajchman M.A., Clarke B.J.;
 RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CIRCULATES IN THE BLOOD IN A ZYMOGEN FORM. FACTOR VII IS
 CC CONVERTED TO FACTOR VIIA BY FACTOR XIa, FACTOR XIIa, FACTOR IXa, OR
 CC THROMBIN BY MINOR PROTEOLYSES. IN THE PRESENCE OF TISSUE FACTOR
 CC AND CALCIUM IONS, FACTOR VIIa THEN CONVERTS FACTOR X TO FACTOR Xa
 CC BY LIMITED PROTEOLYSIS. FACTOR VIIa WILL ALSO CONVERT FACTOR IX TO
 CC FACTOR IXa IN THE PRESENCE OF TISSUE FACTOR AND CALCIUM (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to
 CC form factor Xa.
 CC -1- SUBUNIT: HETERODIMER OF A LIGHT CHAIN AND A HEAVY CHAIN LINKED
 CC BY A DISULFIDE BOND (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: PLASMA.

CC -1- PTH: THE VITAMIN K-DEPENDENT, ENZYMAIC CARBOXYLATION OF SOME
 CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
 CC CALCIUM (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTRAINS 2 EGF-LIKE DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U77477; AAB37326.1; -
 DR HSSP: P08709; IFAK.
 DR MEROPS: S01.215; -
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF-2.
 DR InterPro: IPR001881; EGF-Ca.
 DR InterPro: IPR002383; GLA_blood.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR InterPro: IPR000294; Vitk_dep_GLA.
 DR Pfam: PF00008; EGF; 2.
 DR Pfam: PF00089; trypsin; 1.
 DR Pfam: PF00594; gla; 1.
 DR PRINTS: PRO0722; CHYMOTRYPSIN.
 DR PRINTS: PRO0001; GLABLOOD.
 DR SMART: SM00179; EGF_CA; 1.
 DR SMART: SM00001; EGF_like; 1.
 DR SMART: SM00069; GLA; 1.
 DR SMART: SM00020; Tryp_SPC; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS01187; EGF_CA; 1.
 DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR Hydroxylase; Serine protease; Blood coagulation; Zymogen; Glycoprotein;
 KW Liver; Plasma; Vitamin K; Calcium-binding; Gamma-carboxyglutamic acid;
 KW EGF-like domain; Repeat; Signal; Hydroxylation.
 FT SIGNAL 1 21
 FT PROPEP 22 39
 FT CHAIN 40 191
 FT CHAIN 192 444
 FT DOMAIN 45 74
 FT DOMAIN 85 121
 FT DOMAIN 126 167
 FT DOMAIN 192 444
 FT SITE 191 192
 FT
 FT ACT_SITE 232 232
 FT ACT_SITE 281 281
 FT ACT_SITE 383 383
 FT BINDING 377 377
 FT BINDING 56 61
 FT DISULFID 89 100
 FT DISULFID 94 109
 FT DISULFID 111 120
 FT DISULFID 130 141
 FT DISULFID 137 151
 FT DISULFID 153 166
 FT DISULFID 174 301
 FT DISULFID 198 203
 FT DISULFID 217 233
 FT DISULFID 349 368
 FT DISULFID 379 407
 FT MOD_RES 45 45
 FT MOD_RES 46 46
 GAMMA-CARBOXYGLUTAMIC ACID.
 GAMMA-CARBOXYGLUTAMIC ACID.

FT MOD_RES 53 53
 FT MOD_RES 55 55
 FT MOD_RES 58 58
 FT MOD_RES 59 59
 FT MOD_RES 64 64
 FT MOD_RES 65 65
 FT MOD_RES 68 68
 FT MOD_RES 74 74
 FT MOD_RES 102 102
 FT MOD_RES 211 211
 FT CARBOHYD 242 242
 FT CARBOHYD 306 306
 SQ SEQUENCE 444 AA; 49011 MW; 0481ABC4FE5427F8 CRC64;

Query Match 8.4%; Score 67.5; DB 1; Length 444;
 Best Local Similarity 27.4%; Pred. No. 17;
 Matches 26; Conservative 11; Mismatches 39; Indels 19; Gaps 3;
 QY 24 ICAGKSSNEIRICDRH-----GGGQY-----SAQRSQRPFGVDILCSAGSTV 66
 Db 108 ICFCIADEFGRCNCEKKNKDQLICMYENGGEQYCSDHVGSORSCRCHEGYTLPLPGVSC- 166
 QY 67 YAPFTGMIVGQEKPYONKNAINNGVRIISGRCVCK 101
 Db 167 -TPVDYPCGKVPALKERKGSNPGRIIVGKVCCK 200

Search completed: March 25, 2003, 17:20:06
 Job time : 9.5 secs

Q9DFJ1 ID Q9DFJ1 PRELIMINARY; PRT; 156 AA.
 AC Q9DFJ1;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE Chemotaxin (Fragment).
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 NX NCBI_TaxId=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Bayne C.J., Gerwick L., Fujiki K., Nakao M., Yano T.;
 RT "Immune-relevant (including acute phase) genes identified in the
 RT livers of rainbow trout, *Oncorhynchus mykiss*, by means of suppression
 RT subtractive hybridization.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF271114; AAG28030.1;
 FT NON_TER 156
 SO SEQUENCE 156 AA; 17087 MW; B19067665FDA70DA CRC64;

Query Match 46.5%; Score 375; DB 13; Length 156;
 Best Local Similarity 47.3%; Pred. No. 3.4e-32;
 Matches 69; Conservative 29; Mismatches 46; Indels 2; Gaps 1;

QY 7 LLLAGLSTRALAGPMANICAGKSNERTCDRHGCGQYSAQRPHGVDILCSAGSY 66
 Db 10 VLLIVLSECEMYKFFQQLCDNSNRRRTGQWGGHNGASGRGHOGLDVCNDGAVY 69
 QY 67 YAPFTGMIVGOEKRYON--FNAINGVIRISGRFCVKKMFYIKPIYKGIKGEKIGTL 124
 Db 70 YAPFVKLNGKVIYVTPDKKAALINDGINTLSEGLCFKLFYKPKDKYSSGVKKGQRIGTML 129

QY 125 PLQKYPGIGIOSHVHENCSSDPTAY 150
 Db 130 TMOGVPGITSHVQMCDSKDEPTKF 155

RESULT 3
 Q21241 ID Q21241 PRELIMINARY; PRT; 472 AA.
 AC Q21241;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Hypothetical 53.3 kDa protein.
 GN K05F1.5.
 OS *Caenorhabditis elegans*.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdillida; Rhabdilloidea;
 OC Rhabdillidae; Peloderinae; *Caenorhabditis*.
 NX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode *C. elegans*: a platform for
 RT investigating biology. The *C. elegans* Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Wohldmann P.;
 RT "The sequence of *C. elegans* cosmid K05F1.";
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U29377; AAA68720.2;
 KW Hypothetical protein.
 SO SEQUENCE 472 AA; 53330 MW; D7A8D966A7190C2A CRC64;

Query Match 18.2%; Score 146.5; DB 5; Length 472;
 Best Local Similarity 27.1%; Pred. No. 3.1e-07;
 Matches 45; Conservative 28; Mismatches 62; Indels 31; Gaps 9;

QY 7 LLLAGL-ISTALGPANITAGKSNERTC-DRHG-CGQYSAQR-SQRPHGVDILCSA 62
 Db 146 LILIGLIGGQTLGCKKVCNNKNDFFRCPTDSCGNHTERTSGEIIDGVDRCHL 205

QY 63 GSTVYAPFTGMIVGOEKPY---QNKNAINNGVIRISG---RGFCVKKMFYIKPIYKGP 115
 Db 206 GEPITAPRIGEMVFW-RPYGGRKRSKADGVRLEGTOQMGYAAVHSSVKLSFGGHE 264

QY 116 KKEKIGTLPL---QKVPGIOSHV-----HIENC 142
 Db 265 AGDEIGELNRYCFNDRGQNDVEPHEIRLYKEGRLDPTHTLQNC 310

RESULT 4
 001719 ID 001719 PRELIMINARY; PRT; 689 AA.
 AC 001719;
 DT 01-JUL-1997 (TREMblrel. 04, Created)
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Myb-related transcription factor (Fragment).
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinolideae; Echinodermata; Echinacea; Echinolideae; Strongylocentrotidae;
 OC Strongylocentrotus.
 NX NCBI_TaxId=7668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9808684; PubMed=9428408;
 RA Coffman J.A., Kirchhamer C.V., Harrington M.G., Davidson E.H.;
 RT "SPMYb functions as an intramolecular repressor to regulate spatial
 RT expression of *Cyrlia* in sea urchin embryos.";
 RL Development 124:4717-4727(1997).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 3 MYB-LIKE DOMAINS.
 DR EMBL; U96090; AAC47807.1; -
 DR HSSP; P06876; IMB.
 DR InterPro; IPR001005; Myb_DNA_binding.
 DR Pfam; PF00249; MYB_DNA_binding.
 DR SMART; SM00395; SANT; 3.
 DR PROSITE; PS00337; MYB_1; UNKNOWN_3.
 DR PROSITE; PS00334; MYB_2; 3.
 DR PROSITE; PS50090; MYB_3; 3.
 KW DNA-binding; Nuclear protein.
 FT NON_TER 1
 SO SEQUENCE 689 AA; 77241 MW; 0FA5A8D7AF8766C6 CRC64;

Query Match 10.7%; Score 86.5; DB 5; Length 689;
 Best Local Similarity 21.4%; Pred. No. 1.2;
 Matches 36; Conservative 22; Mismatches 39; Indels 71; Gaps 9;

QY 17 LAGPMANIC--AGKSSNEIR---TCDR--HCGQYSAQR-----SORPHQ 54
 Db 163 LGRNWAETAKLLPGRFDNAIKHNMNSTMKRKVETCNPTPTPKTTKTPHYTHNDQKP-- 220

QY 55 GVDILCSAGSYVYAPFTGMIVGOEKPYQNKNAINNGVIRISGRFCVKKMFYIKPIYKGP 114
 Db 221 -----CSSSSKITYP-----DSDFMANNSIDALMRGQGVY----- 253

QY 115 KKEKIGTLPL-----QKVPGL-----QSHVHENCSSD 147
 Db 254 -----VKTLYPMGHDTLGDDEGSKYKPGVKTQKMLIMNCGEISP 296

RESULT 5
Q96ZEB PRELIMINARY; PRT; 272 AA.
AC Q96ZEB: 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hypothetical protein S71885.
GN S71885.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
NCBI_TaxID=111955;
RN NCB1
RP SEQUENCE FROM N.A.
RC STRAIN-JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankal A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kuchida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagis M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT *Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AF000988; BAB66977.1; "
DR InterPro: IPR000515; BPD.transp.
DR PROSITE: PS00402; BPD.TRANSF.INN.MEMBR; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 272 AA; 30871 MW; 810C97413773BF05 CRC64;

Query Match 10.5%; Score 85; DB 17; Length 272;
Best Local Similarity 26.9%; Pred. No. 0.6;
Matches 28; Conservative 21; Mismatches 43; Indels 12; Gaps 5;

QY 54 QGVDIICASGVYAPFTGMT-----VGDEKPYQNKAINNGVRISGRCVQKMYIK 106
DB 23 KAIDSSPDQENFSPSGIIEKIKRPNKY-TKIDYDVVMYINSGKRIVLHVE 81
QY 107 PIYKKG-PIKKGKLTLLPLQKY-PIQSHVHIENCSSDP 148
DB 82 PYEEGSEIKEGEKIKR--LESPYAGDFKHAHIEGITFEKPS 123

RESULT 6
Q9PTR2 PRELIMINARY; PRT; 498 AA.
AC Q9PTR2: 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Cytochrome P450 2P2.
GN CYP2P2.
OS Fundulus heteroclitus (Killifish) (Mummichog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Fundulidae; Fundulus.
NCBI_TaxID=8078;
RN NCB1
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Oleksiak M.F., Stegeman J.J.;
RT *Diversity of vertebrate cytochrome P450 2 family genes: Cloning of
RT three new genes and phylogenetic analysis of CYP2 gene subfamilies in
RT Fundulus heteroclitus.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN NCB1
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Oleksiak M.F., Wu S., Parker C., Zeldin D.C., Stegeman J.J.;
RT "Cloning and expression of a novel teleost cytochrome P450, CYP2P3;

RT conservation of arachidonic acid epoxigenase/19-Hydroxylase.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF117342; AAF21999.1; "
DR HSSP: P00179; 1PT6.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 498 AA; 57340 MW; 1FF4ED84BBD610C9 CRC64;

Query Match 10.0%; Score 81; DB 13; Length 498;
Best Local Similarity 21.7%; Pred. No. 3.2;
Matches 40; Conservative 29; Mismatches 57; Indels 58; Gaps 10;

QY 1 MFTKALLIAGLISFALGPW-----ANTAGKSSNEFRCTDRHCCGYSAQS 49
DB 117 VFKKKGIVMSN-----GYPMKVQRRFALHALNFGIGKRTME-----RYIQEC 160
QY 50 Q-----RPHQGVDIICASGVYAPFTGMTVGDEKPYQNK-----NAINNGVR 92
DB 161 QYLNVEVVDQGGKPPSGQTLLINNAVSNIT---ICCLVFGNRFEDDKRYHTLIDMNNELR 217
QY 93 ISGRGFCVKMFIKPIK--GPIKKGKLTLLPLQKYPIQISHV--HIENCSSDP 147
DB 218 LQG-GFWQVYVNMFPSPVKMLPGPHKK-----IFIHLQKTIIDFLRIKREHNDLPSSP 271
QY 148 TAYL 151
DB 272 RDTY 275

RESULT 7
Q9PGX9 PRELIMINARY; PRT; 417 AA.
AC Q9PGX9: 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical protein Xf0167.
GN Xf0167.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
NCBI_TaxID=2371;
RN NCB1
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Agencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Farto J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Medeira A.M.B.N., Medeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montello-Vitello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*."
DR Nature 406:151-159(2000).
DR EMBL; AE003870; AAF82980.1; -
DR InterPro; IPR002886; Peptidase_M37.
DR Pfam; PF01551; Peptidase_M37.1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 417 AA; 46554 MW; 5483743CD61061D CRC64;

Query Match 10.0%; Score 80.5; DB 16; Length 417;
Best Local Similarity 25.0%; Pred. No. 3;
Matches 29; Conservative 24; Mismatches 42; Indels 21; Gaps 7;

QY 34 RUCDHGCGQYSAQRSPHQVDILCSAGSTVYAPFTGMI--VGQEKPYONKNAINNGV 91
DB 249 RUSFTFGLRRHPLTGMWRMHKGDYAAATGTPIMACDARVOFTQQRGTGVNVIINH-- 306
QY 92 RISGRGFCVKKMFYIKPIKYGPIKKGKEL--GTLPLQKVPYGIQS----HVHIE 140
DB 307 ---GKGY--RTLYGHMSRF--GKIKAGCKINGTVIG---YVGMGLATGPHLHYE 352

RESULT 8

Q98088 PRELIMINARY; PRT; 750 AA.
AC Q98088;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE LIPOPROTEIN.
GN MYPU_4780.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
CX NCBI_TaxID=2107;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed-11353084;
RA Chambaud I., Hellig R., Ferris S., Barbe V., Samson D., Gallison F.,
RA Moszer I., Dydevig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis."
RL Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL; AL445564; CAC13651.1; -
DR Mypulist; MYPU_4780; -
KW Complete proteome.
SQ SEQUENCE 750 AA; 86325 MW; 37F06FE161602D5 CRC64;

Query Match 9.8%; Score 79; DB 16; Length 750;
Best Local Similarity 22.8%; Pred. No. 8.5;
Matches 42; Conservative 25; Mismatches 55; Indels 62; Gaps 8;

QY 12 LISTALAGPWANICAGKSSNEIRTCDR--HCGQYSAQRSPHQVDILCSAGSTVYAP 69
DB 456 IQSPIYSGWMLD-----DRSIVTSKASDKNTKDIHLGEDILLIQNKKEVIAP 502
QY 70 FTGMIYGO--EKPYQ-----NKNAINGVRIISGRGFCVKKMF-- 103
DB 503 FDGKITASYAPSPYAYAGLVITLVLEMKDLVQIDQSVIDNOLAETDRITYAMHNL 562
QY 104 -----YIKPIKKG-----PIKKGKELTLLPLQKVPYGIQSIVHIE--NCD 143
DB 563 PSYLENYGLVEOSSTALEITPAPKTYAKGEVIG--LVGEFKNGGMPVHIEVSLG 621
QY 144 SSDP 147
DB 622 STNP 625

RESULT 9

Q9UF93 PRELIMINARY; PRT; 651 AA.
AC Q9UF93;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Hypothetical 70.4 kDa protein (fragment).
GN DKF2P566N1047.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Otenwelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wleemann S.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL13109; CAB61414.1; -
DR InterPro; IPR002965; P_rich_extensn.
DR Pfam; PF01424; R3H; 1.
DR PRINTS; PR01217; PRICHEXTENSU.
DR SMART; SM00393; R3H; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 651 AA; 70424 MW; 4A8DF18AB50AD04C CRC64;

Query Match 9.4%; Score 76; DB 4; Length 651;
Best Local Similarity 31.4%; Pred. No. 15;
Matches 27; Conservative 14; Mismatches 25; Indels 20; Gaps 5;

QY 43 QYSAQRSPHQVDILCSAGSTVYAP-----FTGMIYGOEKPYONKNAINNGVRI 93
DB 500 QYNARSQGMQ-----AAQAGCYQVPLSQGQFGRLI-GVQCPQSQNVINNGQSTV 552
QY 94 SGRGFCVKKMFYIKPIKYGPIKKGKEL--GTLPLQKVPYGIQS----HVHIE 140
DB 553 QS---VWVSPTTMSSTVYPMPTGSO 574

RESULT 10

Q9AKK4 PRELIMINARY; PRT; 284 AA.
AC Q9AKK4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 32.4 kDa protein.
OS Rickettsia montana.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsia.
CX NCBI_TaxID=33991;

RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21219194; PubMed-11319266;
RA Andersson J.O., Andersson S.G.E.;
RT "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes."
RL Mol. Biol. Evol. 18:829-839(2001).
DR EMBL; AJ293330; CAC33652.1; -
DR InterPro; IPR002886; Peptidase_M37.
DR Pfam; PF01551; Peptidase_M37.1.
KW Hypothetical protein.
SQ SEQUENCE 284 AA; 32423 MW; 6708E9F3A54EF1F5 CRC64;

Query Match 9.4%; Score 75.5; DB 2; Length 284;
Best Local Similarity 23.6%; Pred. No. 6.5;
Matches 26; Conservative 20; Mismatches 35; Indels 29; Gaps 6;

QY 44 YSAQRSP--HOGVDILCSAGSTVYAPFTGMI--VGQEKPYONKNAINNGVRIISGRGFCV 100
DB 171 YKKRRKKCFHSGIDILQAKKAAPITAAAGTIVAKARAPDIGNFVEIKH-----GKKFTY 225

QY 101 KMYIKPIKYK--GPIKKGKGLTLLPLQKVPICQ-----SHWHE 140
 Db 226 KYAHKMSVKEGNKIKRQ-----FIGIOGSGNATGEHLHFE 264

RESULT 11

Q9IED5 PRELIMINARY; PRT; 538 AA.
 AC 09IED5;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE Env polypeptide (Fragment).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 NCBI_TaxId=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YBF51;
 RA Roques P., Robertson D., Damond F., Souquiere S., Mauciere P.,
 RT "HIV-1 group O phylogenetic analysis of C2-gp41 region."
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ133072; CAB96233.1;
 DR InterPro: IPR000328; Env-GP41.
 DR Pfam: PF00516; GP120.1.
 DR Pfam: PF00517; GP41.1.
 KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.
 FT NON_TER 1
 FT SEQUENCE 538 AA; 60777 MW; B3C9E66A233FEFD CRC64;

Query Match 9.4%; Score 75.5; DB 15; Length 538;
 Best Local Similarity 26.9%; Pred. No. 14;
 Matches 36; Conservative 13; Mismatches 42; Indels 43; Gaps 6;

QY 25 CAGSSN--EIRTCRHGGGQYSAORSORPHGVDLCSAGSYVAP-----FT 71
 Db 175 CEGSCKNKRENT-----NNNTQIPCKLKQYVRSMKSGSLYAPIRGNLCRSMT 227
 QY 72 GMIVGQEKPYONKNAINGVRISG-----RGFCVMEFYIKPIKYGPI--- 114
 Db 228 GMILQDLKPW-NRNDTNTFRPIGGMKMDIWTELLRKVRIKPFVAPRIARPVIGS 286
 QY 115 -----KKGKGLT 123
 Db 287 GTHREKRAVGLGL 300

RESULT 12

Q9HS60 PRELIMINARY; PRT; 240 AA.
 AC 09HS60;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Vng0391C.
 GN VNG0391C.
 OS Halobacterium sp. (strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Halobacterium.
 NCBI_TaxId=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shroana J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,

RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Edhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RT "Genome sequence of Halobacterium species NRC-1."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 DR EMBL: AE004996; AAC18948.1;
 DR InterPro: IPR002761; DDF71.
 DR Pfam: PF01902; DDF71.1.
 DR TIGRFAMs: TIGR00290; Mj0570_dom.1.
 KW Complete proteome.

SO SEQUENCE 240 AA; 25855 MW; FCAE7C983BCEFE72A CRC64;

Query Match 9.3%; Score 75; DB 17; Length 240;
 Best Local Similarity 23.4%; Pred. No. 6.1;
 Matches 37; Conservative 23; Mismatches 56; Indels 42; Gaps 8;

QY 9 LAGLSTLACGPWANICAKSSMEIT-----CDRRGCGYSARSGRPHGVNLLCSA 62
 Db 98 VADTLDGSLAG-----VIAGAVESDFQTSRMEGMDRLGCELYAPLMQPPRLAAMLA 153
 QY 63 GSTVAPFTGMIV-----GOEKPYQKN-----AINN--GVRISSGFCVMEFY 105
 Db 154 G-----FEIRITQVAAYGLDKSMGLRDLRALAALALNDEYGVHMLGEGEETLVT 207
 QY 106 KPIKYPKIKGKGLTLLPLQKVPICQSHVH-ENC 142
 Db 208 DAPHMDRPIR-----LDYETVMAGDRGHIEETAC 237

RESULT 13

Q8RD77 PRELIMINARY; PRT; 270 AA.
 AC 08RD77;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Membrane proteins related to metalloendopeptidases.
 GN NLPD OR TTE0169.
 OS Thermococcus aerobacter tengcongensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Thermococcobacteriales; Thermococcobacteriaceae; Thermococcobacter.
 NCBI_TaxId=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MBAT / JCM11007;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of T. tengcongensis genome."
 RL Genome Res. 12:689-700(2002).
 DR EMBL: AE012990; AAM23470.1;
 KW Complete proteome.

SO SEQUENCE 270 AA; 30340 MW; 27F1342B983E3094 CRC64;
 Query Match 9.3%; Score 75; DB 16; Length 270;
 Best Local Similarity 26.9%; Pred. No. 7;
 Matches 29; Conservative 14; Mismatches 47; Indels 18; Gaps 4;

QY 53 HQGVNLLCSAGSYVAPFTGMIV-----VGOEKPYQKNNAINGVRISGRGFCVMEFYIK----- 106
 Db 168 HKGIDIAKGLPEVVAASDGISSVKKDPLRGNTVVIKNGI-----WEMVYASLDGN 219
 QY 107 -PIKYPKIKGKGLTLLPLQKVPICQSHVH-----ENCSDSPYAT 150
 Db 220 IKVKEGDKITKGOIGEVGDYAKFELAEAGHLELRENGVPIDPTPY 267

RESULT 14

Q9SUS5 PRELIMINARY; PRT; 444 AA.
 AC 09SUS5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:10:14 ; Search time 59 Seconds
(without alignments)
341.031 Million cell updates/sec

Title: us-09-868-953a-2

Perfect score: 807
Sequence: 1 MFSTKALLAGLISTALGP.....GIQSHVHENCSSDPAYL 151

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|----|-------------|
| 1 | 807 | 100.0 | 151 | 19 | AAW52000 |
| 2 | 802 | 99.4 | 151 | 19 | AAW23840 |
| 3 | 802 | 99.4 | 151 | 21 | AAAB1932 |
| 4 | 796 | 98.6 | 151 | 17 | AAK99271 |
| 5 | 796 | 98.6 | 151 | 22 | AAK99271 |
| 6 | 702 | 87.0 | 151 | 19 | AAW57630 |
| 7 | 645 | 79.9 | 133 | 17 | AAK97815 |
| 8 | 642 | 79.6 | 133 | 14 | AAK32177 |
| 9 | 430 | 53.3 | 326 | 23 | AAE16952 |
| 10 | 294 | 36.4 | 54 | 22 | ABB38145 |

| | | | | | |
|----|------|------|-----|----|----------|
| 11 | 294 | 36.4 | 54 | 22 | ABB23353 |
| 12 | 294 | 36.4 | 54 | 22 | AAW58781 |
| 13 | 294 | 36.4 | 54 | 22 | AAW71294 |
| 14 | 294 | 36.4 | 54 | 22 | AAK18972 |
| 15 | 294 | 36.4 | 54 | 22 | AAK31573 |
| 16 | 294 | 36.4 | 54 | 23 | ABG41094 |
| 17 | 256 | 31.7 | 54 | 17 | AAK99266 |
| 18 | 196 | 24.3 | 41 | 17 | AAK97816 |
| 19 | 132 | 16.4 | 27 | 17 | AAK97829 |
| 20 | 95 | 11.8 | 20 | 17 | AAK99270 |
| 21 | 94 | 11.6 | 18 | 17 | AAK97819 |
| 22 | 90 | 11.2 | 19 | 17 | AAK97818 |
| 23 | 89 | 11.0 | 19 | 17 | AAK97827 |
| 24 | 86 | 10.7 | 18 | 17 | AAK97826 |
| 25 | 82 | 10.2 | 25 | 23 | AAE16951 |
| 26 | 77.5 | 9.6 | 580 | 22 | AAK38996 |
| 27 | 76 | 9.4 | 438 | 22 | ABG02090 |
| 28 | 75 | 9.3 | 309 | 21 | AAK29833 |
| 29 | 75 | 9.3 | 349 | 21 | AAK29832 |
| 30 | 75 | 9.3 | 444 | 21 | AAK29831 |
| 31 | 75 | 9.3 | 444 | 23 | ABK93077 |
| 32 | 74.5 | 9.2 | 263 | 23 | AAK88022 |
| 33 | 74 | 9.2 | 22 | 23 | AAK16949 |
| 34 | 74 | 9.2 | 702 | 7 | AAK60230 |
| 35 | 73 | 9.0 | 14 | 17 | AAK97825 |
| 36 | 73 | 9.0 | 16 | 23 | AAE16950 |
| 37 | 73 | 9.0 | 481 | 19 | AAW83216 |
| 38 | 73 | 9.0 | 481 | 22 | AAE04758 |
| 39 | 72.5 | 9.0 | 168 | 22 | AAK91968 |
| 40 | 72.5 | 9.0 | 355 | 21 | AAK27935 |
| 41 | 72.5 | 9.0 | 370 | 20 | AAK27934 |
| 42 | 72 | 8.9 | 159 | 22 | AAK43029 |
| 43 | 70.5 | 8.7 | 625 | 19 | AAW83200 |
| 44 | 70.5 | 8.7 | 625 | 19 | AAW69958 |
| 45 | 70.5 | 8.7 | 625 | 19 | AAW68294 |

ALIGNMENTS

| | | |
|----------|--|----------------------------|
| RESULT 1 | AAW52000 | standard; Protein; 151 AA. |
| ID | AAW52000 | |
| XX | AAW52000 | |
| AC | AAW52000 | |
| XX | AAW52000 | |
| DT | 15-MAR-2002 | (first entry) |
| XX | Human Mim-1 | homologous protein. |
| DE | Human Mim-1 | homologous protein. |
| XX | Human Mim-1 | homologue. |
| KW | Human; Mim-1 | homologue. |
| XX | Human; Mim-1 | homologue. |
| OS | Homo sapiens. | |
| XX | Homo sapiens. | |
| FH | Key | Location/Qualifiers |
| FT | Misc-difference 1 | /note= "encoded by ATO" |
| FT | Misc-difference 58 | /note= "encoded by ATO" |
| FT | Misc-difference 58 | /note= "encoded by ATO" |
| XX | Key | Location/Qualifiers |
| XX | Key | Location/Qualifiers |
| PD | 13-OCT-1997. | |
| XX | 13-OCT-1997. | |
| PF | 29-MAR-1996; | 96KR-0009150. |
| XX | 29-MAR-1996; | 96KR-0009150. |
| PR | 29-MAR-1996; | 96KR-0009150. |
| XX | 29-MAR-1996; | 96KR-0009150. |
| PA | (KOAD) KOREA INST SCI & TECHNOLOGY. | |
| XX | (KOAD) KOREA INST SCI & TECHNOLOGY. | |
| PI | Choi IS, Chung TH, Kim JH, Lee YH, Lee IA, Choi YG, Song JC; | |
| XX | Lim JS; | |
| XX | Lim JS; | |

Protein #5352 enco
Human brain expres
Human bone marrow
Peptide #5406 enco
Peptide #5610 enco
Human peptide enco
Leukocyte cell-der
Bovine chondromodu
Leukocyte cell-der
Bovine chondromodu
Leukocyte cell-der
Bovine chondromodu
Bovine chondromodu
Bovine chondromodu
Chicken myb induce
Human polypeptide
Novel human diagno
Arabidopsis thailia
Arabidopsis thailia
Herbicidally activ
Mouse IIR protein.
Chicken myb induce
Dihydroxyacetone-S
Bovine chondromodu
Chicken myb induce
Human h-RAB-R. Ho
Human vesicle traf
C glutamincum prote
Arabidopsis thailia
Propionibacterium
Murine osteoclast
Murine NF-kB recep
Murine NF-kB recep

DR WPI: 1998-492656/42.
 DR N-PSDB: ABA03485.
 PT Mim-1 homologous gene derived from human tissue NoAbstract -
 XX
 PS Claim 1: Page 2; 4pp; Korean.
 CC The present invention relates to the human Mim-1 homologous gene. The
 CC present sequence is the Mim-1 protein.
 XX
 SQ Sequence 151 AA:
 Query Match 100.0%; Score 807; DB 19; Length 151;
 Best Local Similarity 100.0%; Pred. No. 2, 2e-83;
 Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MESTALLAGLSTALAGWANTCAGKSSNEIRTCDRHGGQYSAORSQRPHGVDTLC 60
 DB 1 MESTALLAGLSTALAGWANTCAGKSSNEIRTCDRHGGQYSAORSQRPHGVDTLC 60
 QY 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVYKMFYIKPIYKGPICKGKEL 120
 DB 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVYKMFYIKPIYKGPICKGKEL 120
 QY 121 GTLLPLQKVPYGIQSHVHIENCSSDPTAYL 151
 DB 121 GTLLPLQKVPYGIQSHVHIENCSSDPTAYL 151

RESULT 2
 AAM23840
 ID AAM23840 standard; Protein; 151 AA.

XX AAM23840;

DT 22-MAY-1998 (first entry)

DE Human LECT2.

KW Human; leukocyte-derived chemotaxin 2; LECT2; monoclonal antibody;
 immunosassay; diagnosis; hepatitis; liver cirrhosis.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Misc-difference 58 /label= Val, Ile

PN WO9745451-A1.

PD 04-DEC-1997.

PF 26-MAY-1997; 97WO-JP01775.

PR 27-MAY-1996; 96JP-0132160.

PA (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.

PI Arai T;

DR WPI: 1998-032582/03.

DR N-PSDB: AAT76922.

PT Monoclonal antibodies which recognise human leukocyte-derived
 chemotaxin 2 - for diagnostic assay of this chemotaxin in hepatitis
 and other diseases

PS Claim 1; Pages 20-21; 43pp; Japanese.

CC The present sequence is human leukocyte-derived chemotaxin 2

CC (LECT2).

CC Novel monoclonal antibodies recognising human LECT2 are produced by
 CC culture of hybridoma G2A5D7 (FERM P-15638), hybridoma A1G1C6

CC (FERM P-15639), hybridoma 5C5 (FERM P-15640), hybridoma H12D10D6
 CC (FERM P-15641) and/or hybridoma 89F2 (FERM P-16229) obtained by
 CC fusion of spleen cells from mice immunised with human LECT2 as
 CC antigen, with mouse myeloma cells. The antibodies are used as the
 CC 1st (immobilised) antibody in ELISA immunoassay for human LECT2,
 CC for the diagnosis of diseases such as hepatitis and liver
 CC cirrhosis.
 XX
 SQ Sequence 151 AA:
 Query Match 99.4%; Score 802; DB 19; Length 151;
 Best Local Similarity 99.3%; Pred. No. 8e-83;
 Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESTALLAGLSTALAGWANTCAGKSSNEIRTCDRHGGQYSAORSQRPHGVDTLC 60
 DB 1 MESTALLAGLSTALAGWANTCAGKSSNEIRTCDRHGGQYSAORSQRPHGVDTLC 60
 QY 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVYKMFYIKPIYKGPICKGKEL 120
 DB 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVYKMFYIKPIYKGPICKGKEL 120
 QY 121 GTLLPLQKVPYGIQSHVHIENCSSDPTAYL 151
 DB 121 GTLLPLQKVPYGIQSHVHIENCSSDPTAYL 151

RESULT 3

AAB11932
 ID AAB11932 standard; Protein; 151 AA.

XX AAB11932;

DT 20-NOV-2000 (first entry)

DE Human leukocyte-derived chemotaxin 2 (LECT2).

KW Human leukocyte-derived chemotaxin 2; LECT2; bone resorption inhibitor;
 leukocyte activating factor; hypercalcaemia; osteoporosis; bone fracture;
 lumbago; lumbodysnia; hyperparathyroidism; Paget's disease.

KW Homo sapiens.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Protein 19..151 /note= "Mature human LECT2"

FT Misc-difference 56 /label= Val, Ile

PN WO200037093-A1.

PD 29-JUN-2000.

PF 20-DEC-1999; 99WO-JP07152.

PR 22-DEC-1998; 98JP-0363727.

PA (MAIN-) NAT INST INFECTIOUS DISEASES.
 (SUZU/) SUZUKI K.

PI Suzuki K, Yamagoe S, Yamakawa T;

DR WPI: 2000-442530/38.

PT Leukocyte activating protein factor LECT2 or its derivative with bone
 resorption inhibitory activity, useful for efficacious therapy of
 hypercalcaemia, osteoporosis, bone fracture, lumbodysnia,
 hyperparathyroidism and Paget's disease

PS Claim 2; Page 14-15; 20pp; Japanese.

CC This sequence represents human leukocyte-derived chemotaxin 2 (LECT2), a
 CC leukocyte activating factor. The invention relates to the use of LECT2

CC and derivatives thereof as bone resorption inhibitors. The invention also
 CC encompasses a method of screening candidate bone resorption inhibitors
 CC using LECT2, and the inhibitors thus identified. Human LECT2 was isolated
 CC from undifferentiated bone cells, and its ability to inhibit osteoclast-
 CC mediated bone resorption was examined via a pit assay. Human LECT2 caused
 CC 50% inhibition of bone resorption at 1 microgram/ml and 100% inhibition
 CC at 10 micrograms/ml. LECT2 or its derivatives may be used for the
 CC treatment of hypercalcaemia, osteoporosis, bone fractures, lumbar
 CC (lumbodynia), hyperparathyroidism and Paget's disease.

SO Sequence 151 AA;
 Query Match 99.4%; Score 802; DB 21; Length 151;
 Best Local Similarity 99.3%; Pred. No. 8e-83;
 Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MFSTKALLAGLISTALAGPMANICAGKSSNEIRTCDRHGGGQYSAQRSPRHQGVDTLC 60
 Db 1 MFSTKALLAGLISTALAGPMANICAGKSSNEIRTCDRHGGGQYSAQRSPRHQGVDTLC 60
 OY 61 SAGSTVYAPFTGMIVGOEKPYONKNAINGVRISGRGFCVVMFYIKPIKYGPPIKKEKL 120
 Db 61 SAGSTVYAPFTGMIVGOEKPYONKNAINGVRISGRGFCVVMFYIKPIKYGPPIKKEKL 120
 OY 121 GTLLPLOKRVPGIOSHVHIENCSDSDPTAYL 151
 Db 121 GTLLPLOKRVPGIOSHVHIENCSDSDPTAYL 151

RESULT 4
 AAR9271
 ID AAR9271 standard; Protein; 151 AA.

AC AAR9271;
 XX 13-DEC-1996 (first entry)

XX Leukocyte cell-derived chemotaxin LECT2b.
 DE LECT2b; LECT2a; leukocyte cell-derived chemotaxin; cytokine;
 KM neutrophil; chemotaxis; cancer; therapy; diagnosis.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH Misc-difference 42 /note= "residue 42 may also be Ile, owing to a
 FT polymorphism in the nucleotide sequence"

XX EP723016-A2.
 PN 24-JUL-1996.
 XX 27-NOV-1995; 95EP-0402670.
 XX 28-NOV-1994; 94JP-0293233.
 XX (SUZU/) SUZUKI K.
 PA Amiga S, Mizuno S, Suzuki K, Yamakawa Y, Yamagoe S;
 PI WPI; 1996-335477/34.
 DR N-PSDB; AAT35260.

PT New human leukocyte cell-derived chemotaxin cpds - used for the
 PT detection and therapy of cancer and disorders of the cytokine
 PT network
 XX Claim 2; Page 13; 17pp; English.
 XX Novel human leukocyte cell-derived chemotaxin LECT2b (AAR9271) has
 CC deduced from a cDNA clone (AAT35260) derived from

CC phytohaemagglutinin-activated T-cell leukaemia cell line SKW-3.
 CC Recombinant LECT2a can be produced, pref. as a fusion protein, in
 CC transformed host cells, esp. E. coli, yeast, insect cells, CHO,
 CC CV-1, 293, C127, 3T3, L-929, HeLa and SKW-3 cells. LECT2b and
 CC LECT2a (see also AAR9266-70) are useful for the diagnosis, therapy
 CC and prediction of cancer and disorders of the cytokine network.

SO Sequence 151 AA;
 Query Match 98.6%; Score 796; DB 17; Length 151;
 Best Local Similarity 98.0%; Pred. No. 3.8e-82;
 Matches 148; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 MFSTKALLAGLISTALAGPMANICAGKSSNEIRTCDRHGGGQYSAQRSPRHQGVDTLC 60
 Db 1 MFSTKALLAGLISTALAGPMANICAGKSSNEIRTCDRHGGGQYSAQRSPRHQGVDTLC 60
 OY 61 SAGSTVYAPFTGMIVGOEKPYONKNAINGVRISGRGFCVVMFYIKPIKYGPPIKKEKL 120
 Db 61 SAGSTVYAPFTGMIVGOEKPYONKNAINGVRISGRGFCVVMFYIKPIKYGPPIKKEKL 120
 OY 121 GTLLPLOKRVPGIOSHVHIENCSDSDPTAYL 151
 Db 121 GTLLPLOKRVPGIOSHVHIENCSDSDPTAYL 151

RESULT 5
 ABG14781
 ID ABG14781 standard; Protein; 681 AA.

AC ABG14781;
 XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #14772.
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.

XX WO200175067-A2.
 PN 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS78968.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

XX Claim 20; SEQ ID No 45140; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.

CC Sequence 681 AA:

Query Match 91.6%; Score 739; DB 22; Length 681;
 Best Local Similarity 100.0%; Pred. No. 8.7e-75;
 Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ALAGPWANICAGKSSNEIRTCDRHGGCGOYSAGRSQRPBGVDILCSAGSTVYAPFTGMTV 75
 |||||
 DB 546 ALAGPWANICAGKSSNEIRTCDRHGGCGOYSAGRSQRPBGVDILCSAGSTVYAPFTGMTV 605

QY 76 GQEKPYQNKNAINGVRISGRFCVKKMFYIKPIYKGPICKGKLGTLPLQKVPYGIOS 135
 |||||
 DB 606 GQEKPYQNKNAINGVRISGRFCVKKMFYIKPIYKGPICKGKLGTLPLQKVPYGIOS 665

QY 136 HVHIENCDSDDPTAYL 151
 |||||
 DB 666 HVHIENCDSDDPTAYL 681

RESULT 6
 AAM57630
 ID AAM57630 standard; Protein; 151 AA.

AC AAM57630;

XX 27-AUG-1998 (first entry)

XX Chondromodulin-II protien.

XX Cow; chondromodulin-II; Chm-II; osteoclast growth; cell growth factor;

KW cartilage growth; bone fracture healing.

XX Bos taurus.

XX Key Location/Qualifiers

FT Peptide 1..18

FT /note= "signal peptide"

FT Protein 9..151

FT /note= "mature chondromodulin-II"

XX JP10146189-A.

XX 02-JUN-1998.

XX 15-NOV-1996; 96JP-0304942.

XX 15-NOV-1996; 96JP-0304942.

XX (MITU) MITSUBISHI CHEM CORP.

XX WPI; 1998-369864/32.

XX N-PSDB; AAV24457.

XX DNA coding chondromodulin-II protein - stimulates increase of
 PT cartilage, and hence healing of bone fractures

XX Claim 2; Page 14; 17pp; Japanese.

XX This sequence is the bovine chondromodulin-II (Chm-II) protein of
 CC the invention, which has the following properties: (1) it is

CC water-soluble, and comprises a polypeptide with a molecular weight of
 CC 16 kD by SDS-PAGE; (2) it has an activity of growing osteoclast alone or
 CC in the co-presence of another cell growth factor; (3) it has an activity
 CC of promoting differentiating function in cartilage cell; and (4) it has
 CC an activity of growing cartilage cell alone or in the co-presence of
 CC another cell growth factor. The protein stimulates the increase of
 CC cartilage, and hence the healing of bone fractures. The Chm-II can supply
 CC recombinant Chm-II stably.

XX Sequence 151 AA:

Query Match 87.0%; Score 702; DB 19; Length 151;
 Best Local Similarity 86.1%; Pred. No. 1.8e-71;
 Matches 130; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 MFSTKALLAGLSTLAGPWANICAGKSSNEIRTCDRHGGCGOYSAGRSQRPBGVDILC 60
 |||||
 DB 1 MFSTKALLAGLSTLAGPWANICAGKSSNEIRTCDRHGGCGOYSAGRSQRPBGVDILC 60

QY 61 SAGSTVYAPFTGMTVGOEKPYQNKNAINGVRISGRFCVKKMFYIKPIYKGPICKGK 120
 |||||
 DB 61 SAGSTVYAPFTGMTVGOEKPYQNKNAINGVRISGRFCVKKMFYIKPIYKGPICKGK 120

QY 121 GTLLPLQKVPYGIOSSHVHIENCDSDDPTAYL 151
 |||||
 DB 121 GTLLPLQKVPYGIOSSHVHIENCDSDDPTAYL 151

RESULT 7
 AAR97815
 ID AAR97815 standard; protein; 133 AA.

AC AAR97815;

XX 22-AUG-1996 (first entry)

XX Bovine chondromodulin-II.

XX bone metabolism; osteoclast formation; bone resorption; disorder;

KW disease; treatment; chondrocyte growth factor; chondromodulin;

KW fibroblast growth factor.

XX Bos taurus.

XX JP08027020-A.

XX 30-JAN-1996.

XX 19-JUL-1994; 94JP-0166917.

XX 19-JUL-1994; 94JP-0166917.

XX (MITU) MITSUBISHI CHEM CORP.

XX WPI; 1996-173892/18.

XX Agent for promoting osteoclast formation and bone absorption
 PT containing chondrocyte growth factor, e.g. chondromodulin, as active
 PT component

XX Claim 4; Page 6; 6pp; Japanese.

XX Bovine chondromodulin-II has mol. wt. 16 kD and promotes chondrocyte
 CC growth (opt. in the presence of fibroblast growth factor). The
 CC protein is the preferred chondrocyte growth factor for use as active
 CC ingredient in a novel agent for promoting osteoclast formation and
 CC bone resorption. Such an agent is intended for treating metabolic
 CC bone diseases.

XX Sequence 133 AA:

XX Query Match 79.9%; Score 645; DB 17; Length 133;
 CC Best Local Similarity 87.2%; Pred. No. 4.4e-65;

Dy
QY 19 GPMANICAGKSSNEIFCTCDRHGGGQYSAPRSQRPHGVDDILCSAGSYTYPFTGMIVGE 78

Db 1 GPAAITCAGKSSNEIFCTDGHGGGQYYAQRNQKLHQGDVDCSDGSYTYPFTGRIMQE 60

| | | | | |
|----|-----------------------------------|-----------------------------------|-----------------------------------|----|
| QY | 7 | LLIAGISTSLAGP----- | WANIICAGSSNEIITPCBHHGGGYSASQSR--p | 52 |
| | | : | : | |
| Db | 7 | IALLSLVSTFAKOMEVHPPOOGRHMAQIICSGP | PNRIKRCDDYGGCGNIGASAKQSGEK | 66 |
| | | | | |
| 53 | HQGVDLICSAGSTVVAPEFTMIGOEKPYONKNA | INGVRIISGRGFCVKMFEIYIKPYKG | 112Y | |

```

Db      67  HKGVNICTDSIVYAPFWMGSLGRFHNHNGNAIDGVOISGSGCYKLVCIHPIRYHG 126
Oy      113 PIKKGKLGTLPLQKVYVPGIOSHVHIENCDSPTAYL 151
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      127 QIQKGQGLGRMLPMOKVFPGLVSHIHVHNCDSPTAYL 165

RESULT 10
ABB38145
ID      ABB38145 standard; Peptide; 54 AA.
XX
AC      ABB38145;
XX
DT      04-FEB-2002 (first entry)
XX
DE      Peptide #5651 encoded by human foetal liver single exon probe.
XX
KW      Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS      Homo sapiens.
XX
PN      W0200157277-A2.
XX
PD      09-AUG-2001.
XX
PE      30-JAN-2001; 2001WO-US00669.
XX
PR      04-FEB-2000; 2000US-0180312.
XX      26-MAY-2000; 2000US-0207456.
XX      30-JUN-2000; 2000US-0608408.
XX      03-AUG-2000; 2000US-0632366.
XX      21-SEP-2000; 2000US-0234687.
XX      27-SEP-2000; 2000US-0236359.
XX      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI: 2001-483447/52.
XX
PT      Human genome-derived single exon nucleic acid probes useful for
XX      analyzing gene expression in human fetal liver -
XX
PS      Claim 27; SEQ ID NO 30780; 639pp + sequence listing; English.
XX
CC      The invention relates to a single exon nucleic acid probe for
XX      measuring human gene expression in a sample derived from human foetal
XX      liver. The single exon nucleic acid probes may be used for predicting,
XX      measuring and displaying gene expression in samples derived from human
XX      fetal liver. The present sequence is a peptide encoded by a single exon
XX      nucleic acid probe of the invention.
XX
CC      Note: The sequence data for this patent did not form part of the
XX      printed specification, but was obtained in electronic format directly
XX      from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence 54 AA;

Query Match      36.4%; Score 294; DB 22; Length 54;
Best Local Similarity 100.0%; Pred. No. 8.6e-26;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      98 FCVKMFYIRPIKKGPIKKGKLGTLPLQKVYVPGIOSHVHIENCDSPTAYL 151
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1 FCVKMFYIRPIKKGPIKKGKLGTLPLQKVYVPGIOSHVHIENCDSPTAYL 54

```

```

DT      23-JAN-2002 (first entry)
XX
DE      Protein #5352 encoded by probe for measuring heart cell gene expression.
XX
KW      Human; gene expression; heart; microarray; vascular system;
XX      cardiovascular disease; hypertension; cardiac arrhythmia;
XX      congenital heart disease.
XX
OS      Homo sapiens.
XX
PN      W0200157274-A2.
XX
PD      09-AUG-2001.
XX
PE      30-JAN-2001; 2001WO-US00666.
XX
PR      04-FEB-2000; 2000US-0180312.
XX      26-MAY-2000; 2000US-0207456.
XX      30-JUN-2000; 2000US-0608408.
XX      03-AUG-2000; 2000US-0632366.
XX      21-SEP-2000; 2000US-0234687.
XX      27-SEP-2000; 2000US-0236359.
XX      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI: 2001-488899/53.
XX
PT      Single exon nucleic acid probes for analyzing gene expression in human
XX      hearts -
XX
PS      Claim 15; SEQ ID NO 25123; 530pp; English.
XX
CC      The present invention relates to single exon nucleic acid probes for
XX      measuring human gene expression in a sample derived from human heart (see
XX      ABA21535-ABA41305). The present sequence is a protein encoded by one such
XX      probe. The probes may be used for predicting, measuring and displaying
XX      gene expression in samples derived from the human heart via microarrays.
XX      By measuring gene expression, the probes are useful for predicting,
XX      diagnosing, grading, staging, monitoring and prognosing diseases of the
XX      human heart and vascular system e.g. cardiovascular disease,
XX      hypertension, cardiac arrhythmias and congenital heart disease.
XX
CC      Note: The sequence data for this patent did not form part of the printed
XX      specification, but was obtained in electronic format directly from WIPO
XX      at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence 54 AA;

Query Match      36.4%; Score 294; DB 22; Length 54;
Best Local Similarity 100.0%; Pred. No. 8.6e-26;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      98 FCVKMFYIRPIKKGPIKKGKLGTLPLQKVYVPGIOSHVHIENCDSPTAYL 151
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1 FCVKMFYIRPIKKGPIKKGKLGTLPLQKVYVPGIOSHVHIENCDSPTAYL 54

RESULT 12
AAM58781
ID      AAM58781 standard; Protein; 54 AA.
XX
AC      AAM58781;
XX
DT      05-NOV-2001 (first entry)
XX
DE      Human brain expressed single exon probe encoded protein SEQ ID NO: 30886.
XX
KW      Human; brain expressed exon; gene expression analysis; probe;
XX      microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX      epilepsy; cancer.
XX

```

OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 FI
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 30886; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 CC
 SQ Sequence 54 AA:
 Query Match 36.4%; Score 294; DB 22; Length 54;
 Best Local Similarity 100.0%; Pred. No. 8.6e-26;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 98 FCVKNFYIKPIKYGKPIKKEKLTGTLPLQKYVPGIOSHVHIENCSSDPTAYL 151
 DB 1 FCVKNFYIKPIKYGKPIKKEKLTGTLPLQKYVPGIOSHVHIENCSSDPTAYL 54
 RESULT 13
 AAM71294
 ID AAM71294 standard; Protein: 54 AA.
 AC AAM71294;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31600.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukemia; lymphoma; myeloma.
 OS Homo sapiens.
 OS
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 FI
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 31600; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 CC
 SQ Sequence 54 AA:
 Query Match 36.4%; Score 294; DB 22; Length 54;
 Best Local Similarity 100.0%; Pred. No. 8.6e-26;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 98 FCVKNFYIKPIKYGKPIKKEKLTGTLPLQKYVPGIOSHVHIENCSSDPTAYL 151
 DB 1 FCVKNFYIKPIKYGKPIKKEKLTGTLPLQKYVPGIOSHVHIENCSSDPTAYL 54
 RESULT 14
 AAM18972
 ID AAM18972 standard; Protein: 54 AA.
 AC AAM18972;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Peptide #5406 encoded by probe for measuring cervical gene expression.
 XX
 KW Probe: human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 OS Homo sapiens.
 OS
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00670.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 FI
 DR WPI; 2001-488901/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -
 XX
 PS Claim 27; SEQ ID NO 23798; 487pp; English.
 XX
 CC The present invention relates to human single exon nucleic acid probes
 CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded

CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 54 AA:

Query Match 36.4%; Score 294; DB 22; Length 54;
 Best Local Similarity 100.0%; Pred. No. 8.6e-26;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 FCVAMFYIKPIKYGPIKKGKGLGTLPLQKYVPGIOSHVHIENCSSDPTAYL 151
 DB 1 FCVAMFYIKPIKYGPIKKGKGLGTLPLQKYVPGIOSHVHIENCSSDPTAYL 54

RESULT 15

AAM31573
 ID AAM31573 standard; Protein; 54 AA.

XX AAM31573:

AC AAM31573:

DT 17-OCT-2001 (first entry)

DE Peptide #5610 encoded by probe for measuring placental gene expression.

KW Probe: microarray; human; placenta; antenatal diagnosis;

XX genetic disorder.

OS Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human placenta -

XX Claim 27; SEQ ID NO 31842; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENP;

XX see AA131315-AA157546). The present sequence is a peptide encoded by one

XX such probe. The probes are useful for producing a microarray for

XX predicting, measuring and displaying gene expression in samples derived

XX from human placenta. The probes are useful for antenatal diagnosis of

XX human genetic disorders.

XX Sequence 54 AA:

OY 98 FCVAMFYIKPIKYGPIKKGKGLGTLPLQKYVPGIOSHVHIENCSSDPTAYL 151
 DB 1 FCVAMFYIKPIKYGPIKKGKGLGTLPLQKYVPGIOSHVHIENCSSDPTAYL 54

Search completed: March 25, 2003, 17:19:37
 Job time : 60 secs

Query Match 36.4%; Score 294; DB 22; Length 54;

Best Local Similarity 100.0%; Pred. No. 8.6e-26;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:10:14 ; Search time 59 Seconds
(without alignments)
341.031 Million cell updates/sec

Title: US-09-868-953A-1
Perfect score: 804
Sequence: 1 MFSTKALLAGLISTALAGP.....GIQSHVHENCSDSPAYL 151

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 802 | 99.8 | 151 | 19 | AAAS2000 |
| 2 | 802 | 99.8 | 151 | 19 | AAAS2000 |
| 3 | 802 | 99.8 | 151 | 21 | AAAB1932 |
| 4 | 792 | 98.5 | 151 | 17 | AAAB9271 |
| 5 | 734 | 91.3 | 681 | 22 | ABG14781 |
| 6 | 698 | 86.8 | 151 | 19 | AAAS7630 |
| 7 | 641 | 79.7 | 133 | 17 | AAAB97815 |
| 8 | 638 | 79.4 | 133 | 14 | AAAB32177 |
| 9 | 426 | 53.0 | 326 | 23 | AAAB16952 |
| 10 | 294 | 36.6 | 54 | 22 | ABAB38145 |

| | | | | | |
|----|------|------|-----|----|-----------|
| 11 | 294 | 36.6 | 54 | 22 | ABAB23353 |
| 12 | 294 | 36.6 | 54 | 22 | AAAB58781 |
| 13 | 294 | 36.6 | 54 | 22 | AAAB71294 |
| 14 | 294 | 36.6 | 54 | 22 | AAAB18972 |
| 15 | 294 | 36.6 | 54 | 22 | AAAB31573 |
| 16 | 294 | 36.6 | 54 | 23 | ABG41094 |
| 17 | 252 | 31.3 | 54 | 17 | AAAB92666 |
| 18 | 192 | 23.9 | 41 | 17 | AAAB97816 |
| 19 | 132 | 16.4 | 27 | 17 | AAAB97816 |
| 20 | 95 | 11.8 | 20 | 17 | AAAB9270 |
| 21 | 94 | 11.7 | 18 | 17 | AAAB97819 |
| 22 | 89 | 11.1 | 19 | 17 | AAAB97827 |
| 23 | 86 | 10.7 | 18 | 17 | AAAB97826 |
| 24 | 86 | 10.7 | 19 | 17 | AAAB97818 |
| 25 | 78 | 9.7 | 25 | 23 | AAAB16951 |
| 26 | 76 | 9.5 | 438 | 22 | ABG02090 |
| 27 | 75.5 | 9.4 | 580 | 22 | AAAB8996 |
| 28 | 74 | 9.2 | 22 | 23 | AAAB16949 |
| 29 | 74 | 9.2 | 702 | 7 | AAAB60230 |
| 30 | 73 | 9.1 | 14 | 17 | AAAB97825 |
| 31 | 73 | 9.1 | 16 | 23 | AAAB16950 |
| 32 | 73 | 9.1 | 159 | 22 | AAAB43029 |
| 33 | 72.5 | 9.0 | 715 | 20 | AAAB05625 |
| 34 | 71.5 | 8.9 | 706 | 23 | AAAB92784 |
| 35 | 70.5 | 8.8 | 642 | 23 | AAAB97592 |
| 36 | 70.5 | 8.8 | 680 | 22 | AAAB92854 |
| 37 | 70.5 | 8.8 | 680 | 23 | AAAB97564 |
| 38 | 70 | 8.7 | 309 | 21 | AAAB29833 |
| 39 | 70 | 8.7 | 349 | 21 | AAAB29833 |
| 40 | 70 | 8.7 | 444 | 21 | AAAB29831 |
| 41 | 70 | 8.7 | 444 | 23 | AAAB93077 |
| 42 | 69.5 | 8.6 | 168 | 22 | AAAB91968 |
| 43 | 69.5 | 8.6 | 263 | 23 | AAAB80822 |
| 44 | 69 | 8.6 | 14 | 17 | AAAB97824 |
| 45 | 68 | 8.5 | 214 | 22 | AAAB41669 |

ALIGNMENTS

| RESULT 1 | ID | AAAS2000 | standard; Protein; 151 AA. |
|----------|-------------|--|----------------------------|
| XX | XX | AAAS2000; | |
| AC | XX | AAAS2000; | |
| DT | 15-MAR-2002 | (first entry) | |
| DE | Human Mim-1 | homologous protein. | |
| XX | XX | Human: Mim-1 | homologue. |
| KW | XX | Human: Mim-1 | homologue. |
| XX | XX | Human sapiens. | |
| OS | XX | Human sapiens. | |
| XX | XX | Key | Location/Qualifiers |
| FT | FT | Misc-difference 1 | /note= "encoded by ATO" |
| FT | FT | Misc-difference 58 | /note= "encoded by ATO" |
| FT | FT | Misc-difference 58 | /note= "encoded by ATO" |
| XX | XX | KR97065715-A. | |
| XX | XX | 13-OCT-1997. | |
| PD | XX | 13-OCT-1997. | |
| XX | XX | 29-MAR-1996; | 96KR-0009150. |
| PF | XX | 29-MAR-1996; | 96KR-0009150. |
| XX | XX | 29-MAR-1996; | 96KR-0009150. |
| PR | XX | (KOAD) KOREA INST SCI & TECHNOLOGY. | |
| XX | XX | Chol IS, Chung TH, Kim JH, Lee YH, Lee IA, Choi YG, Song JC; | |
| PI | XX | Lim JS; | |
| XX | XX | | |

Protein #5352 enco
Human brain expres
Human bone marrow
Peptide #5406 enco
Peptide #5610 enco
Human peptide enco
Leukocyte cell-der
Bovine chondromodu
Bovine chondromodu
Leukocyte cell-der
Bovine chondromodu
Bovine chondromodu
Bovine chondromodu
Chicken myb induce
Novel human diagno
Human polypeptide
Chicken myb induce
Dihydroxyacetone-s
Bovine chondromodu
Chicken myb induce
Protonibacterium
HIV-1 group O Isol
Herbically activ
Novel human protei
Human protein sequ
Novel human protei
Arabidopsis thalia
Arabidopsis thalia
Herbically activ
C glutamicum prote
Mouse Irf8 protein.
Bovine chondromodu
Protonibacterium

DR WPI: 1998-492656/42.
 DR N-PSDB: ABA03485.
 XX
 XX Mim-1 homologous gene derived from human tissue NCABSTRACT -
 XX
 XX Claim 1; Page 2; 4pp; Korean.
 XX
 CC The present invention relates to the human Mim-1 homologous gene. The
 CC present sequence is the Mim-1 protein.
 XX
 SQ Sequence 151 AA:
 Query Match 99.8%; Score 802; DB 19; Length 151;
 Best Local Similarity 99.3%; Pred. No. 2.5e-87;
 Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MFSTALLAGLISLTALAGPWANICAGKSSNEIRTCRHGGCGYSAQSRPHGVDXLC 60
 DB 1 MFSTALLAGLISLTALAGPWANICAGKSSNEIRTCRHGGCGYSAQSRPHGVDXLC 60
 QY 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVAMFYIKPIKYPKIKGKEL 120
 DB 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVAMFYIKPIKYPKIKGKEL 120
 QY 121 GTLLPLQKVPYPGIQSHVHIENCDSDEPTAYL 151
 DB 121 GTLLPLQKVPYPGIQSHVHIENCDSDEPTAYL 151
 RESULT 2
 AAW23840
 ID AAW23840 standard; Protein; 151 AA.
 XX
 AC AAW23840;
 XX
 DT 22-MAY-1998 (first entry)
 DE Human LECT2.
 XX
 KW Human; leukocyte-derived chemotaxin 2; LECT2; monoclonal antibody;
 KW immunosassay; diagnosis; hepatitis; liver cirrhosis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT MISC-difference 58
 FT /label= Val, Ile
 XX
 FN WO9745451-A1.
 XX
 PD 04-DEC-1997.
 XX
 PE 26-MAY-1997; 97WO-JP01775.
 XX
 PR 27-MAY-1996; 96UP-0132160.
 XX
 PA (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.
 XX
 PI Arai T;
 XX
 DR WPI: 1998-032582/03.
 DR N-PSDB: AAT76922.
 XX
 PT Monoclonal antibodies which recognise human leukocyte-derived
 PT chemotaxin 2 - for diagnostic assay of this chemotaxin in hepatitis
 PT and other diseases
 XX
 XX Claim 1; Pages 20-21; 43pp; Japanese.
 XX
 CC The present sequence is human leukocyte-derived chemotaxin 2
 CC (LECT2).
 CC Novel monoclonal antibodies recognising human LECT2 are produced by
 CC culture of hybridoma G2A5D7 (FERM P-15638), hybridoma A1G1C6

CC (FERM P-15639), hybridoma 5C5 (FERM P-15640), hybridoma H12D10D6
 CC (FERM P-15641) and/or hybridoma 89F2 (FERM P-16229) obtained by
 CC fusion of spleen cells from mice immunised with human LECT2 as
 CC antigen, with mouse myeloma cells. The antibodies are used as the
 CC 1st (immobilised) antibody in ELISA immunoassay for human LECT2,
 CC for the diagnosis of diseases such as hepatitis and liver
 CC cirrhosis.
 XX
 SQ Sequence 151 AA:
 Query Match 99.8%; Score 802; DB 19; Length 151;
 Best Local Similarity 100.0%; Pred. No. 2.5e-87;
 Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MFSTALLAGLISLTALAGPWANICAGKSSNEIRTCRHGGCGYSAQSRPHGVDXLC 60
 DB 1 MFSTALLAGLISLTALAGPWANICAGKSSNEIRTCRHGGCGYSAQSRPHGVDXLC 60
 QY 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVAMFYIKPIKYPKIKGKEL 120
 DB 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVAMFYIKPIKYPKIKGKEL 120
 QY 121 GTLLPLQKVPYPGIQSHVHIENCDSDEPTAYL 151
 DB 121 GTLLPLQKVPYPGIQSHVHIENCDSDEPTAYL 151
 RESULT 3
 AAB11932
 ID AAB11932 standard; Protein; 151 AA.
 XX
 AC AAB11932;
 XX
 DT 20-NOV-2000 (first entry)
 DE Human leukocyte-derived chemotaxin 2 (LECT2).
 XX
 KW Human leukocyte-derived chemotaxin 2; LECT2; bone resorption inhibitor;
 KW leukocyte activating factor; hypercalcaemia; osteoporosis; bone fracture;
 KW lumbago; lumbodysnia; hyperparathyroidism; Paget's disease.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Protein 19..151
 FT /note= "Mature human LECT2"
 FT MISC-difference 56
 FT /label= Val, Ile
 XX
 FN WO200037093-A1.
 XX
 PD 29-JUN-2000.
 XX
 PE 20-DEC-1999; 99WO-JP07152.
 XX
 PR 22-DEC-1998; 98UP-0363727.
 XX
 PA (NAIN-) NAT INST INFECTIOUS DISEASES.
 XX
 PI (SUZU/) SUZUKI K.
 XX
 DR Suzuki K, Yamagoe S, Yamakawa T;
 DR WPI: 2000-442530/38.
 XX
 PT Leukocyte activating protein factor LECT2 or its derivative with bone
 PT resorption inhibitory activity, useful for efficacious therapy of
 PT hypercalcaemia, osteoporosis, bone fracture, lumbodysnia,
 PT hyperparathyroidism and Paget's disease -
 XX
 XX Claim 2; Page 14-15; 20pp; Japanese.
 XX
 CC This sequence represents human leukocyte-derived chemotaxin 2 (LECT2), a
 CC leukocyte activating factor. The invention relates to the use of LECT2

CC and derivatives thereof as bone resorption inhibitors. The invention also
 CC encompasses a method of screening candidate bone resorption inhibitors
 CC using LECT2, and the inhibitors thus identified. Human LECT2 was isolated
 CC from undifferentiated bone cells, and its ability to inhibit osteoclast-
 CC mediated bone resorption was examined via a pit assay. Human LECT2 caused
 CC 50% inhibition of bone resorption at 1 microgram/ml and 100% inhibition
 CC at 10 micrograms/ml. LECT2 or its derivatives may be used for the
 CC treatment of hypercalcaemia, osteoporosis, bone fractures, lumbago
 CC (lumbodysnia), hyperparathyroidism and Paget's disease.

XX Sequence 151 AA:

Query Match 99.8%; Score 802; DB 21; Length 151;
 Best Local Similarity 100.0%; Pred. No. 2.5e-87;
 Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MESTKALLAGLSTALAGPWANICAGKSSNEIRPCDRHGGCGQYSAQRORPHQGVXLC 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MESTKALLAGLSTALAGPWANICAGKSSNEIRPCDRHGGCGQYSAQRORPHQGVXLC 60
 OY 61 SAGSTVYAPFTGMIVGEEKPYQNKNAIINGVRIISGRGFCVKMFYIKPIYKGPPIKKGKEL 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 SAGSTVYAPFTGMIVGEEKPYQNKNAIINGVRIISGRGFCVKMFYIKPIYKGPPIKKGKEL 120
 OY 121 GTLLPLQKYPGQISHVHIENCSSDPTAYL 151
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 GTLLPLQKYPGQISHVHIENCSSDPTAYL 151

RESULT 4

AAR99271 ID AAR99271 standard; Protein; 151 AA.

XX AAR99271:

DT 13-DEC-1996 (first entry)

XX Leukocyte cell-derived chemotaxin LECT2b.

XX LECT2b; LECT2a; leukocyte cell-derived chemotaxin; cytokine;

KW neutrophil; chemotaxis; cancer; therapy; diagnosis.

XX Homo sapiens.

OS Location/Qualifiers

PH Key

FT Misc-difference 42 /note= "residue 42 may also be ile, owing to a

FT polymorphism in the nucleotide sequence"

XX EP723016-A2.

XX 24-JUL-1996.

XX 27-NOV-1995; 95EP-0402670.

XX 28-NOV-1994; 94JP-0293233.

XX (SUZU/) SUZUKI K.

XX Arita S, Mizuno S, Suzuki K, Yamakawa Y, Yamagoe S;

XX WPI: 1996-335477/34.

XX N-PSDB; AAT35260.

XX New human leukocyte cell-derived chemotaxin cpds. - used for the

XX network

XX Claim 2; Page 13; 17pp; English.

XX Novel human leukocyte cell-derived chemotaxin LECT2b (AAR99271) has

CC neutrophil chemotactic activity. Its amino acid sequence was

CC deduced from a cDNA clone (AAT35260) derived from

CC phytohemagglutinin-activated T-cell leukaemia cell line SKW-3.
 CC Recombinant LECT2a can be produced, pref. as a fusion protein, in
 CC transformed host cells, esp. E. coli, yeast, insect cells, CHO,
 CC CV-1, 293, C127, 3T3, L-929, HeLa and SKW-3 cells. LECT2b and
 CC LECT2a (see also AAR9266-70) are useful for the diagnosis, therapy
 CC and prediction of cancer and disorders of the cytokine network.

XX Sequence 151 AA:

Query Match 98.5%; Score 792; DB 17; Length 151;
 Best Local Similarity 98.0%; Pred. No. 4e-86;
 Matches 148; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 MESTKALLAGLSTALAGPWANICAGKSSNEIRPCDRHGGCGQYSAQRORPHQGVXLC 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MESTKALLAGLSTALAGPWANICAGKSSNEIRPCDRHGGCGQYSAQRORPHQGVXLC 60
 OY 61 SAGSTVYAPFTGMIVGEEKPYQNKNAIINGVRIISGRGFCVKMFYIKPIYKGPPIKKGKEL 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 SAGSTVYAPFTGMIVGEEKPYQNKNAIINGVRIISGRGFCVKMFYIKPIYKGPPIKKGKEL 120
 OY 121 GTLLPLQKYPGQISHVHIENCSSDPTAYL 151
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 GTLLPLQKYPGQISHVHIENCSSDPTAYL 151

RESULT 5

ABG14781 ID ABG14781 standard; Protein; 681 AA.

XX ABG14781:

DT 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #14772.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001MO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YF;

XX WPI: 2001-639362/73.

XX N-PSDB; AAS78968.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity

XX Claim 20; SEQ ID No 45140; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 681 AA;

Query Match 91.3%; Score 734; DB 22; Length 681;
Best Local Similarity 99.3%; Pred. No. 2.5e-78;
Matches 135; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 16 ALAGPMANICAGKSSNEIRCDRHGGCGYSQRSORPHGVYDLCASGTVYAPFTGMTV 75
DB 546 ALAGPMANICAGKSSNEIRCDRHGGCGYSQRSORPHGVYDLCASGTVYAPFTGMTV 605
QY 76 GOEKPYONKNAIINGVRIISGRGFCVKMFYIKPIKYPKIKKGEKLTLLPLQKVYPGIOS 135
DB 606 GOEKPYONKNAIINGVRIISGRGFCVKMFYIKPIKYPKIKKGEKLTLLPLQKVYPGIOS 665
QY 136 HVHIENCSSDPTAYL 151
DB 666 HVHIENCSSDPTAYL 681

RESULT 6
AAM57630
ID AAM57630 standard; Protein; 151 AA.

XX AAM57630;

XX 27-AUG-1998 (first entry)

XX Chondromodulin-II protein.

XX Cow; chondromodulin-II; ChM-II; osteoclast growth; cell growth factor;
XX cartilage growth; bone fracture healing.

XX Bos taurus.

XX Key Location/Qualifiers

XX Peptide 1..18

XX Protein /note= "signal peptide"

XX /note= "mature chondromodulin-II"

XX JP10146189-A.

XX 02-JUN-1998.

XX 15-NOV-1996; 96JP-0304942.

XX 15-NOV-1996; 96JP-0304942.

XX (MITU) MITSUBISHI CHEM CORP.

XX WPI: 1998-369864/32.

XX N-PSDB: AAY24457.

XX DNA coding chondromodulin-II protein - stimulates increase of
XX cartilage, and hence healing of bone fractures

XX Claim 2; Page 14; 17pp; Japanese.

XX This sequence is the bovine chondromodulin-II (ChM-II) protein of
CC the invention, which has the following properties: (1) It is

CC water-soluble, and comprises a polypeptide with a molecular weight of
CC 16 kD by SDS-PAGE; (2) it has an activity of growing osteoclast alone or
CC in the co-presence of another cell growth factor; (3) it has an activity
CC of promoting differentiating function in cartilage cell; and (4) it has
CC an activity of growing cartilage cell alone or in the co-presence of
CC another cell growth factor. The protein stimulates the increase of
CC cartilage, and hence the healing of bone fractures. The ChM-II can supply
CC recombinant ChM-II stably.

SO Sequence 151 AA;

Query Match 86.8%; Score 698; DB 19; Length 151;
Best Local Similarity 86.1%; Pred. No. 6.3e-75;
Matches 130; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 MESTKALLAGLSTALAGPMANICAGKSSNEIRCDRHGGCGYSQRSORPHGVYDLC 60
DB 1 MESTKALLAGLSTALAGPMANICAGKSSNEIRCDRHGGCGYSQRSORPHGVYDLC 60
QY 61 SAGSTYVAPFTGMTVGOEKPYONKNAIINGVRIISGRGFCVKMFYIKPIKYPKIKKGEKLT 120
DB 61 SAGSTYVAPFTGMTVGOEKPYONKNAIINGVRIISGRGFCVKMFYIKPIKYPKIKKGEKLT 120
QY 121 GTLLPLQKVYPGIOSHVHIENCSSDPTAYL 151
DB 121 GTLLPLQKVYPGIOSHVHIENCSSDPTAYL 151

RESULT 7
AAR97815
ID AAR97815 standard; protein; 133 AA.

XX AAR97815;

XX 22-AUG-1996 (first entry)

XX Bovine chondromodulin-II.

XX bone metabolism; osteoclast formation; bone resorption; disorder;
XX disease; treatment; chondrocyte growth factor; chondromodulin;
XX fibroblast growth factor.

XX Bos taurus.

XX JP08027020-A.

XX 30-JAN-1996.

XX 19-JUL-1994; 94JP-0166917.

XX 19-JUL-1994; 94JP-0166917.

XX (MITU) MITSUBISHI CHEM CORP.

XX WPI: 1996-173892/18.

XX Agent for promoting osteoclast formation and bone absorption

XX containing chondrocyte growth factor, e.g. chondromodulin, as active

XX component

XX Claim 4; Page 6; 6pp; Japanese.

XX Bovine chondromodulin-II has mol. wt. 16 kD and promotes chondrocyte

XX growth (opt. in the presence of fibroblast growth factor). The

XX protein is the preferred chondrocyte growth factor for use as active

XX ingredient in a novel agent for promoting osteoclast formation and

XX bone resorption. Such an agent is intended for treating metabolic

XX bone diseases.

XX Sequence 133 AA;

XX Query Match 79.7%; Score 641; DB 17; Length 133;
XX Best Local Similarity 87.2%; Pred. No. 3.3e-68;

OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 XX
 PD 09-AUG-2001.
 XX
 PE 30-JAN-2001; 2001WO-US000667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 30886; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 XX
 SQ Sequence 54 AA:
 Query Match 36.6%; Score 294; DB 22; Length 54;
 Best Local Similarity 100.0%; Pred. No. 2.2e-27;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 98 FCVKMFYIKPIKYKGPICKGKGLTLLPLQKVPGIOSHVIENCSSDPAYL 151
 DB 1 FCVKMFYIKPIKYKGPICKGKGLTLLPLQKVPGIOSHVIENCSSDPAYL 54
 RESULT 13
 AAM71294
 ID AAM71294 standard; Protein; 54 AA.
 XX
 AC AAM71294;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31600.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 30-JAN-2001; 2001WO-US000668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 31600; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 XX
 SQ Sequence 54 AA:
 Query Match 36.6%; Score 294; DB 22; Length 54;
 Best Local Similarity 100.0%; Pred. No. 2.2e-27;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 98 FCVKMFYIKPIKYKGPICKGKGLTLLPLQKVPGIOSHVIENCSSDPAYL 151
 DB 1 FCVKMFYIKPIKYKGPICKGKGLTLLPLQKVPGIOSHVIENCSSDPAYL 54
 RESULT 14
 AAM18972
 ID AAM18972 standard; Protein; 54 AA.
 XX
 AC AAM18972;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Peptide #5406 encoded by probe for measuring cervical gene expression.
 XX
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 30-JAN-2001; 2001WO-US000670.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488901/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -
 XX
 PS Claim 27; SEQ ID NO 23798; 487pp; English.
 XX
 CC The present invention relates to human single exon nucleic acid probes
 CC (SNP: see AAI10068-AA128459). The present sequence is a peptide encoded

CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 54 AA;

Query Match 36.6%; Score 294; DB 22; Length 54;
 Best Local Similarity 100.0%; Pred. No. 2.2e-27;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 98 FCVKMFYIKPIKKGPIKKEKLTLLPLQKVPYGIOSHVHIENCSSDPTAYL 151
 Db 1 FCVKMFYIKPIKKGPIKKEKLTLLPLQKVPYGIOSHVHIENCSSDPTAYL 54

RESULT 15

AAM31573
 ID AAM31573 standard; Protein; 54 AA.

XX AC AAM31573;

XX DT 17-OCT-2001 (first entry)

XX DE Peptide #5610 encoded by probe for measuring placental gene expression.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;
 XX genetic disorder.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00663.

XX PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488897/53.

XX PT Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human placenta -

XX PS Claim 27; SEQ ID No 31842; 654bp; English.

XX CC The present invention relates to single exon nucleic acid probes (SENP;
 CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.

XX SO Sequence 54 AA;

Query Match 36.6%; Score 294; DB 22; Length 54;
 Best Local Similarity 100.0%; Pred. No. 2.2e-27;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 98 FCVKMFYIKPIKKGPIKKEKLTLLPLQKVPYGIOSHVHIENCSSDPTAYL 151
 Db 1 FCVKMFYIKPIKKGPIKKEKLTLLPLQKVPYGIOSHVHIENCSSDPTAYL 54

Search completed: March 25, 2003, 17:19:36
 Job time : 62 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:17:45 ; Search time 19 Seconds

(without alignments)
233,835 Million cell updates/sec

Title: US-09-868-953A-1

Perfect score: 804
Sequence: 1 MFSTKALLLAGLISTALGP.....GIQSHVHINCDSDDPYAYL 151

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

- 1: Issued_Patents_AA.*
- 2: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 802 | 99.8 | 151 | 2 | US-08-563-148E-6 |
| 2 | 802 | 99.8 | 151 | 4 | US-09-194-139-1 |
| 3 | 641 | 79.7 | 133 | 1 | US-07-924-753-15 |
| 4 | 417 | 51.9 | 98 | 4 | US-09-194-139-9 |
| 5 | 252 | 31.3 | 54 | 2 | US-08-563-148E-1 |
| 6 | 192 | 23.9 | 41 | 1 | US-07-924-753-1 |
| 7 | 132 | 16.4 | 27 | 1 | US-07-924-753-14 |
| 8 | 95 | 11.8 | 20 | 2 | US-08-563-148E-5 |
| 9 | 94 | 11.7 | 18 | 1 | US-07-924-753-4 |
| 10 | 89 | 11.1 | 19 | 1 | US-07-924-753-12 |
| 11 | 86 | 10.7 | 18 | 1 | US-07-924-753-11 |
| 12 | 86 | 10.7 | 19 | 1 | US-07-924-753-3 |
| 13 | 74 | 9.2 | 702 | 6 | 5240838-15 |
| 14 | 73 | 9.2 | 14 | 1 | US-07-924-753-10 |
| 15 | 69 | 8.6 | 14 | 1 | US-07-924-753-9 |
| 16 | 67.5 | 8.4 | 625 | 3 | US-08-996-139-15 |
| 17 | 67.5 | 8.4 | 625 | 4 | US-08-995-659-15 |
| 18 | 67.5 | 8.4 | 625 | 4 | US-09-215-649A-15 |
| 19 | 67.5 | 8.4 | 625 | 4 | US-09-577-780-15 |
| 20 | 67 | 8.3 | 10182 | 4 | US-09-134-001C-3159 |
| 21 | 66.5 | 8.3 | 2285 | 4 | US-09-308-375-2 |
| 22 | 65.5 | 8.1 | 1241 | 4 | US-09-040-774-2 |
| 23 | 64 | 8.0 | 428 | 3 | US-09-118-319-5 |
| 24 | 63.5 | 7.9 | 2595 | 4 | US-09-036-987A-2 |
| 25 | 63.5 | 7.9 | 2595 | 4 | US-09-370-700-2 |
| 26 | 63 | 7.8 | 14 | 4 | US-09-194-139-10 |
| 27 | 63 | 7.8 | 2004 | 1 | US-08-375-709-15 |

| | | | | | | |
|----|------|-----|------|---|-------------------|-------------------|
| 28 | 63 | 7.8 | 2004 | 1 | US-08-752-929-15 | Sequence 15, Appl |
| 29 | 63 | 7.8 | 2004 | 4 | US-09-090-793-9 | Sequence 9, Appl |
| 30 | 62.5 | 7.8 | 406 | 1 | US-08-293-778-24 | Sequence 24, Appl |
| 31 | 62.5 | 7.8 | 406 | 1 | US-08-295-411-5 | Sequence 5, Appl |
| 32 | 62.5 | 7.8 | 406 | 2 | US-08-955-471-5 | Sequence 5, Appl |
| 33 | 62.5 | 7.8 | 406 | 5 | PCT-US92-10242-5 | Sequence 5, Appl |
| 34 | 62.5 | 7.8 | 433 | 3 | US-08-926-842B-61 | Sequence 61, Appl |
| 35 | 62.5 | 7.8 | 444 | 1 | US-08-475-845-2 | Sequence 2, Appl |
| 36 | 62.5 | 7.8 | 444 | 2 | US-08-327-690-2 | Sequence 2, Appl |
| 37 | 62.5 | 7.8 | 444 | 2 | US-08-660-289-2 | Sequence 2, Appl |
| 38 | 62.5 | 7.8 | 444 | 2 | US-08-537-807-2 | Sequence 2, Appl |
| 39 | 62.5 | 7.8 | 444 | 2 | US-08-871-003-2 | Sequence 2, Appl |
| 40 | 62.5 | 7.8 | 444 | 3 | US-08-464-233-2 | Sequence 2, Appl |
| 41 | 62.5 | 7.8 | 444 | 4 | US-09-189-607-2 | Sequence 2, Appl |
| 42 | 62.5 | 7.8 | 444 | 4 | US-09-378-907-2 | Sequence 2, Appl |
| 43 | 62.5 | 7.8 | 444 | 5 | PCT-US94-05779-2 | Sequence 2, Appl |
| 44 | 62.5 | 7.8 | 466 | 1 | US-07-882-202A-4 | Sequence 4, Appl |
| 45 | 62.5 | 7.8 | 466 | 1 | US-08-021-615A-4 | Sequence 4, Appl |

ALIGNMENTS

RESULT 1
US-08-563-148E-6
Sequence 6, Application US/08563148E
Patent No. 5929224
GENERAL INFORMATION:
APPLICANT: Kazuo SUZUKI et al.
TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSSEE: 663-2, Shligi
STREET: Misaki-machi
CITY: Isumi-gun
STATE: Chiba-ken
COUNTRY: JAPAN
ZIP: 299-45
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS v.5
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/563,148E
FILING DATE: No. 5929224ember 27, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 293233/1994
FILING DATE: 28-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: C. Bruce Hamburg
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-5230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)966-2340
TELEFAX: (212)953-7733
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 151 amino acids
TYPE: amino acid
STRANDEDNESS: single strand
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: intermediate fragment
ORIGINAL SOURCE:
ORGANISM: human
TISSUE TYPE: T-cell derived leukemia cells
PUBLICATION INFORMATION:
AUTHORS: Kazuo SUZUKI et al.
TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 to 151
US-08-563-148E-6

Query Match 99.8%; Score 802; DB 2; Length 151;
Best Local Similarity 99.3%; Pred. No. 1,1e-89;
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESTKALLAGLISALAGWANICAGKSSNEIRTCDRHGGCGQYSAQRSPRHQGVDXLC 60
DB 1 MESTKALLAGLISALAGWANICAGKSSNEIRTCDRHGGCGQYSAQRSPRHQGVDXLC 60
QY 61 SAGSTVYAPFTGMIVGQEKPYQNKAINNGVIRISGRGFCVKMFYIKPIKYGKIKGKGL 120
DB 61 SAGSTVYAPFTGMIVGQEKPYQNKAINNGVIRISGRGFCVKMFYIKPIKYGKIKGKGL 120
QY 121 GTLLPLQKYYPGIQSHVHIENCSSDPTAYL 151
DB 121 GTLLPLQKYYPGIQSHVHIENCSSDPTAYL 151

RESULT 2

US-09-194-139-1
Sequence 1, Application US/09194139A
Patent No. 6306608
GENERAL INFORMATION:
APPLICANT: Arai, Takao
TITLE OF INVENTION: ANTI-HUMAN LECT2 ANTIBODY, CELLS PRODUCING THE SAME,
TITLE OF INVENTION: AND METHOD AND KIT FOR ASSAYING THE SAME
FILE REFERENCE: US application 09/194,139
CURRENT APPLICATION NUMBER: US/09/194,139A
EARLIER FILING DATE: 1998-11-25
EARLIER APPLICATION NUMBER: JP 8-132160
EARLIER FILING DATE: 1996-05-27
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 151
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: residue 58 is either Valine or Isoleucine
US-09-194-139-1

Query Match 99.8%; Score 802; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 1,1e-89;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESTKALLAGLISALAGWANICAGKSSNEIRTCDRHGGCGQYSAQRSPRHQGVDXLC 60
DB 1 MESTKALLAGLISALAGWANICAGKSSNEIRTCDRHGGCGQYSAQRSPRHQGVDXLC 60
QY 61 SAGSTVYAPFTGMIVGQEKPYQNKAINNGVIRISGRGFCVKMFYIKPIKYGKIKGKGL 120
DB 61 SAGSTVYAPFTGMIVGQEKPYQNKAINNGVIRISGRGFCVKMFYIKPIKYGKIKGKGL 120
QY 121 GTLLPLQKYYPGIQSHVHIENCSSDPTAYL 151
DB 121 GTLLPLQKYYPGIQSHVHIENCSSDPTAYL 151

RESULT 3

US-07-924-753-15
Sequence 15, Application US/07924753
Patent No. 5270303
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 133 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:

NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:

US-07-924-753-15

Query Match 79.7%; Score 641; DB 1; Length 133;
Best Local Similarity 87.2%; Pred. No. 3.2e-70;
Matches 116; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 19 GPWANCAGKSSNEIRTCDRHGGCGQYSAQRSPRHQGVDXLCSAGSTVYAPFTGMIVGQE 78
DB 1 GPWANCAGKSSNEIRTCDRHGGCGQYSAQRSPRHQGVDXLCSAGSTVYAPFTGMIVGQE 78

QY 79 KPYONKNAIINNGVIRISGRGFCVAKMEFYIKPIKYKGPRIKKEKLTLLPLQKYVPGIOSHVN 138
 DB 61 KPYONKNAIINNGVIRISGRGFCVAKMEFYIKPIKYKGIKKKKEKLTLLPLQKYVPGIOSHVN 120
 QY 139 IENCDDSDPTAVL 151
 DB 121 IENCDDSDPTAVL 133

RESULT 4
 US-09-194-139-9
 ; Sequence 9, Application US/09194139A
 ; Patent No. 6306608
 ; GENERAL INFORMATION:
 ; APPLICANT: Arai, Takao
 ; TITLE OF INVENTION: ANTI-HUMAN TECT2 ANTIBODY, CELLS PRODUCING THE SAME,
 ; TITLE OF INVENTION: AND METHOD AND KIT FOR ASSAYING THE SAME
 ; FILE REFERENCE: US Application 09/194,139
 ; CURRENT APPLICATION NUMBER: US/09/194,139A
 ; EARLIER FILING DATE: 1998-11-25
 ; EARLIER FILING DATE: 1996-05-27
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 98
 ; TYPE: PRT
 ; ORGANISM: Bovine
 ; US-09-194-139-9

Query Match

Best Local Similarity 51.9%; Score 417; DB 4; Length 98;
 Matches 84; Conservative 6; Mismatches 8; Indels 32; Gaps 4;

QY 19 GPMANICAGKSSNEIRTCRHHGCGOYTAORNOKLGHOVDVLCDSGSTVYAPPTG 78
 DB 1 GPMANICAGKSSNEIRTCRHHGCGOYTAORNOKLGHOVDVLCDSGSTVYAPPTG-IMGOE 59

QY 79 KPYONKNAIINNGVIRISGRGFCVAKMEFYIKPIKYKGPRIKKEKLTLLPLQKYVPGIOSHVN 138
 DB 60 KPYONKNAIINNGVIRISGRGFCVAKMEFYIKPIKYKGIKKKKEKLTLLPLQKYVPGIOSHVN 120

QY 139 IENCDDSDPT 148
 DB 89 IENCDDSDPT 98

RESULT 5
 US-08-563-148E-1
 ; Sequence 1, Application US/08563148E
 ; Patent No. 5929224
 ; GENERAL INFORMATION:
 ; APPLICANT: Kazuo SUZUKI et al.
 ; TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESS: 663-2, Shijiri
 ; STREET: Misaki-machi
 ; CITY: Isumi-gun
 ; STATE: Chiba-ken
 ; COUNTRY: JAPAN
 ; ZIP: 299-45
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: MS-DOS V.5
 ; SOFTWARE: Word Perfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/563,148E
 ; FILING DATE: No. 5929224ember 27, 1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 293233/1994
 ; FILING DATE: 28-NOV-1994

ATTORNEY/AGENT INFORMATION:
 NAME: C. Bruce Hamburg
 REGISTRATION NUMBER: 22,389
 REFERENCE/DOCKET NUMBER: F-5230
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)986-2340
 TELEFAX: (212)953-7733
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 54 amino acids
 TYPE: amino acid
 STRANDEDNESS: single strand
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: N-terminal fragment
 ORIGINAL SOURCE:
 ORGANISM: human
 TISSUE TYPE: T-cell derived leukemia cells
 PUBLICATION INFORMATION:
 AUTHORS: Kazuo SUZUKI et al.
 TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
 RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 to 54
 US-08-563-148E-1

Query Match
 Best Local Similarity 31.3%; Score 252; DB 2; Length 54;
 Matches 45; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 19 GPMANICAGKSSNEIRTCRHHGCGOYTAORNOKLGHOVDVLCDSGSTVYAPPTG 72
 DB 1 GPMANICAGKSSNEIRTCRHHGCGOYTAORNOKLGHOVDVLCDSGSTVYAPPTG 54

RESULT 6

US-07-924-753-1
 ; Sequence 1, Application US/07924753
 ; Patent No. 5270303
 ; GENERAL INFORMATION:
 ; APPLICANT: Fujio SUZUKI et al.
 ; TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESS: Wenderoth, Lind & Ponack
 ; STREET: 805 Fifteenth Street, N.W., #700
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: DisplayWrite
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/924,753
 ; FILING DATE: 19920804
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Warren M. Cheek, Jr.
 ; REGISTRATION NUMBER: 33,367
 ; REFERENCE/DOCKET NUMBER:
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-371-8850
 ; TELEFAX:
 ; TELEPHONE:
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 41 amino acids
 ; TYPE: AMINO ACID
 ; STRANDEDNESS: single

```

? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHETICAL:
? ANTI-SENSE:
? FRAGMENT TYPE: N-terminal fragment
? ORIGINAL SOURCE:
? ORGANISM: bovine
? STRAIN:
? INDIVIDUAL ISOLATE:
? DEVELOPMENTAL STAGE:
? HAPLOTYPE:
? TISSUE TYPE: fetal cartilage
? CELL TYPE:
? ORGANELLE:
? IMMEDIATE SOURCE:
? LIBRARY:
? CLONE:
? POSITION IN GENOME:
? CHROMOSOME/SEGMENT:
? MAP POSITION:
? UNITS:
? FEATURE:
? NAME/KEY:
? LOCATION:
? IDENTIFICATION METHOD:
? OTHER INFORMATION:
? PUBLICATION INFORMATION:
? AUTHORS:
? TITLE:
? JOURNAL:
? VOLUME:
? ISSUE:
? PAGES:
? DATE:
? DOCUMENT NUMBER:
? FILING DATE:
? PUBLICATION DATE:
? RELEVANT RESIDUES IN SEQ ID NO:
? US-07-924-753-1

Query Match      23.9%; Score 192; DB 1; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.3e-16;
Matches 34; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 19 GPMANICAGKSSNEIRTCDRHGGCGVSAQSRSPHGVYDYL 59
    ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GPMATICAGKSSNEIRTCDSHGCGGYTAQRNOKLHGVDYL 41

RESULT 7
US-07-924-753-14
; Sequence 14, Application US/07924753
; Patent No. 5270303
; GENERAL INFORMATION:
; APPLICANT: Fujio SUZUKI et al.
; TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,753
; FILING DATE: 19920804
```

```

? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Warren M. Cheek, Jr.
? REGISTRATION NUMBER: 33,367
? REFERENCE/DOCKET NUMBER:
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-371-8850
? TELEFAX:
? INFORMATION FOR SEQ ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 27 amino acids
? TYPE: AMINO ACID
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHETICAL:
? ANTI-SENSE:
? FRAGMENT TYPE: internal fragment
? ORIGINAL SOURCE:
? ORGANISM: bovine
? STRAIN:
? INDIVIDUAL ISOLATE:
? DEVELOPMENTAL STAGE:
? HAPLOTYPE:
? TISSUE TYPE: fetal cartilage
? CELL TYPE:
? CELL LINE:
? ORGANELLE:
? IMMEDIATE SOURCE:
? LIBRARY:
? CLONE:
? POSITION IN GENOME:
? CHROMOSOME/SEGMENT:
? MAP POSITION:
? UNITS:
? FEATURE:
? NAME/KEY:
? LOCATION:
? IDENTIFICATION METHOD:
? OTHER INFORMATION:
? PUBLICATION INFORMATION:
? AUTHORS:
? TITLE:
? JOURNAL:
? VOLUME:
? ISSUE:
? PAGES:
? DATE:
? DOCUMENT NUMBER:
? FILING DATE:
? PUBLICATION DATE:
? RELEVANT RESIDUES IN SEQ ID NO:
? US-07-924-753-14

Query Match      16.4%; Score 132; DB 1; Length 27;
Best Local Similarity 92.6%; Pred. No. 2.6e-09;
Matches 25; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 111 KPIKKGKLTLLPLQKYVPGIQSHV 137
    || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 KGIKKGKLTLLPLQKYVPGIQSHI 27

RESULT 8
US-08-563-148E-5
; Sequence 5, Application US/08563148E
; Patent No. 5929224
; GENERAL INFORMATION:
; APPLICANT: Kazuo SUZUKI et al.
```

```

; TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES: 12
; ADDRESSEE: 663-2, Shitagi
; STREET: Misaki-machi
; CITY: Isumi-gun
; STATE: Chiba-ken
; COUNTRY: JAPAN
; ZIP: 299-45
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS v. 5
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/563,148E
; FILING DATE: NO. 5929224ember 27, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 293233/1994
; FILING DATE: 28-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5230
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single strand
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: intermediate fragment
; ORIGINAL SOURCE:
; ORGANISM: human
; TISSUE TYPE: T-cell derived leukemia cells
; PUBLICATION INFORMATION:
; AUTHORS: Kazuo SUZUKI et al.
; TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 to 20
; US-08-563-148E-5

Query Match 11.8%; Score 95; DB 2; Length 20;
Best Local Similarity 85.0%; Pred. No. 5.4e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 129 VYPGIQSHVHIEHCSSDPT 148
      |||||||:|||||
Db 1 VYPGIQSHVHIEHCSSDPT 20

RESULT 9
US-07-924-753-4
; Sequence 4, Application US/07924753
; Patent No. 5270303
; GENERAL INFORMATION:
; APPLICANT: Fujio SUZUKI et al.
; TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; City: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Displaywrite

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,753
; FILING DATE: 19920804
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE: internal fragment
; ORIGINAL SOURCE:
; ORGANISM: bovine
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE: fetal cartilage
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-924-753-4

Query Match 11.7%; Score 94; DB 1; Length 18;
Best Local Similarity 88.9%; Pred. No. 6.2e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 129 VYPGIQSHVHIEHCSSD 146
      |||||||:|||||
Db 1 VYPGIQSHVHIEHCSSD 18

RESULT 10
US-07-924-753-12
; Sequence 12, Application US/07924753

```

Patent No. 5270303
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
MOLECULE TYPE: linear
TOPOLOGY: linear
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:

FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-12
Query Match 11.1%; Score 89; DB 1; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00027;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 68 APTGMIWGEEKPYKNKA 86
DB 1 APTGKIMGEEKPYKNKKA 19
RESULT 11
US-07-924-753-11
Sequence 11, Application US/07924753
Patent No. 5270303
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
MOLECULE TYPE: protein
TOPOLOGY: linear
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:

```

; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-11

```

```

Query Match      10.7%; Score 86; DB 1; Length 18;
Best Local Similarity 94.1%; Pred. No. 0.00058;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 79 KPYONKNAINNGVRISG 95
Db 1 KPYONKNAINNGVRISG 17

```

```

RESULT 12
; Sequence 3, Application US/07924753
; Patent No. 5270303
; GENERAL INFORMATION:
; APPLICANT: Fujio SUZUKI et al.
; TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Displaywrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,753
; FILING DATE: 19920804
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

```

; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE: Internal fragment
; ORIGINAL SOURCE:
; ORGANISM: bovine
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE: fetal cartilage
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-3

```

```

Query Match      10.7%; Score 86; DB 1; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.00062;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 53 HGVVDVLCDSAGSTVYAPF 70
Db 2 HGVVDVLCDSAGSTVYAPF 19

```

```

RESULT 13
; Patent No. 5240838
; APPLICANT: LEDERER, ADRIANUS M.; MAAT, JAN; VERRIPS, CORNELIS P.;
; T. VISSER, CHRISTIAN; JANOWICZ, ZBIGNIEW A.; HOLLNBERG, CORNELIS P.
; TITLE OF INVENTION: REGULATORY SEQUENCES OF ALCOHOL OXIDASE
; (MOX) AND DIHYDROXYACETONESYNTASE (DAS) OF HANSBNUA POLYMORPHA
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 702
; TYPE:
; STRANDEDNESS:
; TOPOLOGY:
; MOLECULE TYPE:
5240838-15

```

```

Query Match      9.2%; Score 74; DB 6; Length 702;
Best Local Similarity 24.1%; Pred. No. 2.3;
Matches 33; Conservative 22; Mismatches 48; Indels 34; Gaps 7;
QY 11 GLISTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAQRSQPRHGVVDVLCDSAGSTVYAPF 70

```

Db 525 GLFOKAVELPFSSILS-LSRNEVL-----CYLASRAQRNRNAGYILLEDAENAEVOI 575
QY 71 TGMTVGOEKPYONKNAINNVRISGRF-----CYKMFYIRPIKKGPI--KKGEL 120
Db 576 IG-VGAEMEFPADKA-----KITGRKFRVRVLSIPCTRLPDEOSIGYRSVLRDRGROV 628
QY 121 GTLLPLQKVPKIQSHV 137
Db 629 PTVV-----VDAHV 637

RESULT 14
US-07-924-753-10
; Sequence 10, Application US/07924753
; Patent No. 5270303
; GENERAL INFORMATION:
; APPLICANT: FUJIO SUZUKI et al.
; TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,753
; FILING DATE: 19920804
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE: internal fragment
; ORIGINAL SOURCE:
; ORGANISM: bovine
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE: fetal cartilage
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:

UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-10

Query Match 9.1%; Score 73; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 KIGTLPLQKVPK 132
Db 1 KIGTLPLQKVPK 14

RESULT 15
US-07-924-753-9
; Sequence 9, Application US/07924753
; Patent No. 5270303
; GENERAL INFORMATION:
; APPLICANT: FUJIO SUZUKI et al.
; TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,753
; FILING DATE: 19920804
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:
; ANTI-SENSE:

```

; FRAGMENT TYPE: internal fragment
; ORIGINAL SOURCE:
; ORGANISM: bovine
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE: fetal cartilage
; CELL TYPE:
; CELL LINE:
; ORGANELLAE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-9

Query Match      8.6%; Score 69; DB 1; Length 14;
Best Local Similarity 92.9%; Pred. No. 0.048;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 103 FYIKPIKYKSP1KK 116
Db 1 FYIKPIKYKSP1KK 14

```

Search completed: March 25, 2003, 17:23:00
Job time : 28 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:21:15 ; Search time 12 Seconds
(without alignments)
672.733 Million cell updates/sec

Title: US-09-868-953a-1

Perfect score: 804

Sequence: 1 MFSTKALLAGLSTALGP.....GIQSHVHTIENDSSDPTAYL 151

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PC1_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PC105_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 540 | 67.2 | 113 | US-09-864-570-9 | Sequence 9, Appl1 |
| 2 | 540 | 67.2 | 113 | US-10-157-457-9 | Sequence 9, Appl1 |
| 3 | 426 | 53.0 | 326 | US-09-884-570-8 | Sequence 8, Appl1 |
| 4 | 426 | 53.0 | 326 | US-10-157-457-8 | Sequence 8, Appl1 |
| 5 | 348 | 43.3 | 115 | US-09-884-570-11 | Sequence 11, Appl1 |
| 6 | 348 | 43.3 | 115 | US-10-157-457-11 | Sequence 11, Appl1 |
| 7 | 343 | 42.7 | 115 | US-09-884-570-10 | Sequence 10, Appl1 |
| 8 | 343 | 42.7 | 115 | US-10-157-457-10 | Sequence 10, Appl1 |
| 9 | 294 | 36.6 | 54 | US-09-864-761-38651 | Sequence 38651, A |
| 10 | 78 | 9.7 | 25 | US-09-884-570-7 | Sequence 7, Appl1 |
| 11 | 78 | 9.7 | 25 | US-10-157-457-7 | Sequence 7, Appl1 |
| 12 | 74 | 9.2 | 22 | US-09-884-570-5 | Sequence 5, Appl1 |
| 13 | 74 | 9.2 | 22 | US-10-157-457-5 | Sequence 5, Appl1 |
| 14 | 73 | 9.1 | 16 | US-09-884-570-6 | Sequence 6, Appl1 |
| 15 | 73 | 9.1 | 16 | US-10-157-457-6 | Sequence 6, Appl1 |
| 16 | 69.5 | 8.6 | 168 | US-09-738-626-5722 | Sequence 5722, Ap |
| 17 | 67.5 | 8.4 | 625 | US-09-877-650-15 | Sequence 15, Appl1 |
| 18 | 67.5 | 8.4 | 625 | US-10-166-232A-2 | Sequence 2, Appl1 |
| 19 | 67.5 | 8.4 | 625 | US-09-871-856-15 | Sequence 15, Appl1 |

| | | | | | |
|----|------|-----|------|---------------------|---------------------|
| 20 | 66.5 | 8.3 | 2285 | US-09-932-183A-2 | Sequence 2, Appl1 |
| 21 | 65.5 | 8.1 | 502 | US-09-738-626-4722 | Sequence 4722, Ap |
| 22 | 65.5 | 8.1 | 1241 | US-09-759-632-2 | Sequence 2, Appl1 |
| 23 | 64.5 | 8.0 | 197 | US-09-924-340-66 | Sequence 66, Appl1 |
| 24 | 64.5 | 8.0 | 197 | US-09-992-600A-66 | Sequence 66, Appl1 |
| 25 | 64.5 | 8.0 | 197 | US-09-731-872-276 | Sequence 276, Appl1 |
| 26 | 64 | 8.0 | 598 | US-09-854-845-41 | Sequence 41, Appl1 |
| 27 | 64 | 8.0 | 697 | US-09-854-845-25 | Sequence 25, Appl1 |
| 28 | 63.5 | 7.9 | 151 | US-09-925-301-874 | Sequence 874, Appl1 |
| 29 | 63.5 | 7.9 | 615 | US-09-768-779A-2 | Sequence 2, Appl1 |
| 30 | 63 | 7.8 | 332 | US-09-925-300-1064 | Sequence 1064, Ap |
| 31 | 63 | 7.8 | 1615 | US-09-931-375A-2 | Sequence 2, Appl1 |
| 32 | 62.5 | 7.8 | 375 | US-09-864-761-36336 | Sequence 36336, A |
| 33 | 62.5 | 7.8 | 406 | US-10-109-498-1 | Sequence 1, Appl1 |
| 34 | 62.5 | 7.8 | 603 | US-09-854-845-39 | Sequence 39, Appl1 |
| 35 | 62.5 | 7.8 | 691 | US-10-101-921-4 | Sequence 4, Appl1 |
| 36 | 62.5 | 7.8 | 691 | US-09-925-731-2 | Sequence 2, Appl1 |
| 37 | 62.5 | 7.8 | 702 | US-09-854-845-23 | Sequence 23, Appl1 |
| 38 | 62.5 | 7.8 | 723 | US-10-133-912-18 | Sequence 18, Appl1 |
| 39 | 61.5 | 7.6 | 444 | US-09-840-795-8 | Sequence 8, Appl1 |
| 40 | 61.5 | 7.6 | 492 | US-10-087-667-6 | Sequence 6, Appl1 |
| 41 | 61.5 | 7.6 | 655 | US-09-978-295A-64 | Sequence 64, Appl1 |
| 42 | 61.5 | 7.6 | 655 | US-09-756-854-2 | Sequence 2, Appl1 |
| 43 | 61.5 | 7.6 | 655 | US-09-978-697-64 | Sequence 64, Appl1 |
| 44 | 61.5 | 7.6 | 655 | US-10-041-574-2 | Sequence 2, Appl1 |
| 45 | 61.5 | 7.6 | 655 | US-09-978-192A-64 | Sequence 64, Appl1 |

ALIGNMENTS

RESULT 1

US-09-884-570-9

Sequence 9, Application US/09884570

Patent No. US20020114779A1

GENERAL INFORMATION:

APPLICANT: Williams, John P.

APPLICANT: McDonald, Jay M.

APPLICANT: McKenna, Margaret A.

TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof

FILE REFERENCE: D6294

CURRENT APPLICATION NUMBER: US/09/884,570

CURRENT FILING DATE: 2001-06-19

PRIOR APPLICATION NUMBER: US 60/212,271

PRIOR FILING DATE: 2000-06-19

NUMBER OF SEQ ID NOS: 11

SEQ ID NO 9

LENGTH: 113

TYPE: PRT

ORGANISM: unknown

FEATURE:

LOCATION: 38..165

OTHER INFORMATION: amino acid sequence of region of Lect2

OTHER INFORMATION: protein homologous to repeat sequences of N

OTHER INFORMATION: and C terminal ends of the mim-I protein

US-09-884-570-9

Query Match

Best Local Similarity 67.2%; Score 540; DB 10; Length 113;

Matches 97; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 39 HGGGQYSAORQSRHGGVYDLCASAGSTVYAPFTGMVGGKPYONKNAINGVRSRGCF 98

DB 1 YGGGQYSAORQSRHGGVYDLCASAGSTVYAPFTGMVGGKPYONKNAINGVRSRGCF 60

QY 99 CVKMFYIKPIKYGKPIKGEKIGTLLPDKVYPIGSHVHTIENDSSDPTAYL 151

DB 61 CVKMFYIKPIKYGKPIKGEKIGTLLPDKVYPIGSHVHTIENDSSDPTAYL 113

RESULT 2

US-10-157-457-9

Sequence 9, Application US/10157457

Patent No. US20020150568A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294CIP
CURRENT FILING DATE: 2002-05-29
PRIORITY FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 9
LENGTH: 113
TYPE: PRT
ORGANISM: unknown
FEATURE:
LOCATION: 38..165
OTHER INFORMATION: amino acid sequence of region of Lect2
OTHER INFORMATION: protein homologous to repeat sequences of N
OTHER INFORMATION: and C terminal ends of the mim-1 protein
US-10-157-457-9

Query Match 67.2%; Score 540; DB 12; Length 113;
Best Local Similarity 85.8%; Pred. No. 4.7e-55;
Matches 97; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

OY 39 HGGGYSAORSORPHGVXLCASGTVYAPFTGMIVGOEKPYONKNAIINNGVRISGRGF 98
DB 1 YGCGYSAORTORHNGVYVAFPTGKIYGOEKPYRKNAINDSIRLSGRGF 60
DB 61 CVKMEYIKPKYKGIKKEKLTLLPLQKYPGQISHVHENCDSPTAYL 113
OY 99 CVKMEYIKPKYKGIKKEKLTLLPLQKYPGQISHVHENCDSPTAYL 151
DB 61 CVKMEYIKPKYKGIKKEKLTLLPLQKYPGQISHVHENCDSPTAYL 113

RESULT 3
US-09-884-570-8
Sequence 8, Application US/09884570
Patent No. US20020114779A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294
CURRENT FILING DATE: 2001-06-19
PRIORITY FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 8
LENGTH: 326
TYPE: PRT
ORGANISM: Gallus gallus
FEATURE:
OTHER INFORMATION: amino acid sequence of mim-1 protein
US-09-884-570-8

Query Match 53.0%; Score 426; DB 10; Length 326;
Best Local Similarity 50.3%; Pred. No. 2.6e-41;
Matches 80; Conservative 25; Mismatches 40; Indels 14; Gaps 2;

OY 7 LLAGLSTALAGP-----WANTCAGKSSNEIRPCDRHCGCOYSAORSOR--P 52
DB 7 IALLSLVSTAFAROMVEHPPOQGRHWAQICSGNPFNRIRRCDCRGCGNYGASRGKGR 66
OY 53 HGVVDXLCASGTVYAPFTGMIVGOEKPYONKNAIINNGVRISGRGFVCMFYIKPKYKG 112
DB 67 HKGVDTICTDGSIVYAPFPGQLSPTRFFHNGNAIDGVOISGSGYCVKLVCIHPIRYHG 126
OY 113 PIKKEKLTLLPLQKYPGQISHVHENCDSPTAYL 151
DB 113 PIKKEKLTLLPLQKYPGQISHVHENCDSPTAYL 151

DB 127 QIQKGQOLGRMLPMQKVPFGIVSHIHVENCDSPTHL 165

RESULT 4
US-10-157-457-8
Sequence 8, Application US/10157457
Patent No. US20020150568A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294CIP
CURRENT FILING DATE: 2002-05-29
PRIORITY FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 8
LENGTH: 326
TYPE: PRT
ORGANISM: Gallus gallus
FEATURE:
OTHER INFORMATION: amino acid sequence of mim-1 protein
US-10-157-457-8

Query Match 53.0%; Score 426; DB 12; Length 326;
Best Local Similarity 50.3%; Pred. No. 2.6e-41;
Matches 80; Conservative 25; Mismatches 40; Indels 14; Gaps 2;

OY 7 LLAGLSTALAGP-----WANTCAGKSSNEIRPCDRHCGCOYSAORSOR--P 52
DB 7 IALLSLVSTAFAROMVEHPPOQGRHWAQICSGNPFNRIRRCDCRGCGNYGASRGKGR 66
OY 53 HGVVDXLCASGTVYAPFTGMIVGOEKPYONKNAIINNGVRISGRGFVCMFYIKPKYKG 112
DB 67 HKGVDTICTDGSIVYAPFPGQLSPTRFFHNGNAIDGVOISGSGYCVKLVCIHPIRYHG 126
OY 113 PIKKEKLTLLPLQKYPGQISHVHENCDSPTAYL 151
DB 127 QIQKGQOLGRMLPMQKVPFGIVSHIHVENCDSPTHL 165

RESULT 5
US-09-884-570-11
Sequence 11, Application US/09884570
Patent No. US20020114779A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294
CURRENT FILING DATE: 2001-06-19
PRIORITY FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 11
LENGTH: 115
TYPE: PRT
ORGANISM: Gallus gallus
FEATURE:
LOCATION: 200..315
OTHER INFORMATION: amino acid sequence of repeat sequence of C
OTHER INFORMATION: terminal end of mim-1 protein homologous
OTHER INFORMATION: to Lect2 protein
US-09-884-570-11

Query Match 43.3%; Score 348; DB 10; Length 115;
Best Local Similarity 54.4%; Pred. No. 7.3e-33;
Matches 62; Conservative 24; Mismatches 26; Indels 2; Gaps 1;

PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 38651
LENGTH: 54
TYPE: prt
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004763.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8
OTHER INFORMATION: SWISSPROT HIT: O14960, EVALU8 8.00e-17
US-09-864-761-38651

Query Match 36.6%; Score 294; DB 10; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.9e-27;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 98 FCVKMFYIPKIKKPKKKGKLGTLPLQKVVYPIQSHVHHCNDSSPTAYL 151
Db 1 FCVKMFYIPKIKKPKKKGKLGTLPLQKVVYPIQSHVHHCNDSSPTAYL 54

RESULT 10
US-09-884-570-7
Sequence 7, Application US/09884570
Patent No. US20020114779A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294
CURRENT APPLICATION NUMBER: US/09/884,570
CURRENT FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/212,271
PRIOR FILING DATE: 2000-06-19
NUMBER OF SEQ ID NOS: 11

SEQ ID NO 7
LENGTH: 25
TYPE: prt
ORGANISM: Gallus gallus
FEATURE:
LOCATION: 219..243
OTHER INFORMATION: amino acid sequence tryptic peptide from
OTHER INFORMATION: mim-1 protein
US-09-884-570-7

Query Match 9.7%; Score 78; DB 10; Length 25;
Best Local Similarity 56.0%; Pred. No. 0.017;
Matches 14; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Oy 55 GVDXLCASGTVYAPFGMIVGOEK 79
Db 1 GVDXLCASGTVYAPFGMIVGOEK 25

RESULT 11
US-10-157-457-7
Sequence 7, Application US/10157457
Patent No. US20020150568A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294CIP
CURRENT APPLICATION NUMBER: US/10/157,457
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: US 09/884,570
PRIOR FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 7
LENGTH: 25
TYPE: prt
ORGANISM: Gallus gallus
FEATURE:
LOCATION: 219..243
OTHER INFORMATION: amino acid sequence tryptic peptide from
OTHER INFORMATION: mim-1 protein
US-10-157-457-7

Query Match 9.7%; Score 78; DB 12; Length 25;
Best Local Similarity 56.0%; Pred. No. 0.017;
Matches 14; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Oy 55 GVDXLCASGTVYAPFGMIVGOEK 79
Db 1 GVDXLCASGTVYAPFGMIVGOEK 25

RESULT 12
US-09-884-570-5
Sequence 5, Application US/09884570
Patent No. US20020114779A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294
CURRENT APPLICATION NUMBER: US/09/884,570
CURRENT FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/212,271
PRIOR FILING DATE: 2000-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 5
LENGTH: 22
TYPE: prt
ORGANISM: Gallus gallus
FEATURE:

LOCATION: 94..115
OTHER INFORMATION: amino acid sequence tryptic peptide from
OTHER INFORMATION: mim-1 protein
US-09-884-570-5

Query Match 9.2%; Score 74; DB 10; Length 22;
Best Local Similarity 61.9%; Pred. No. 0.042;
Matches 13; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 81 YONKAINNGVIRISGRFCVK 101
: | | | : | | | : | | | : | | |
Db 2 FHNQNAIDDDGVQISGSGYCVK 22

RESULT 13
US-10-157-457-5
Sequence 5, Application US/10157457
Patent No. US20020150568A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294CIP
CURRENT APPLICATION NUMBER: US/10/157,457
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: US 09/884,570
PRIOR FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 5
LENGTH: 22
TYPE: PRT
ORGANISM: Gallus gallus
FEATURE:
LOCATION: 94..115
OTHER INFORMATION: amino acid sequence tryptic peptide from
OTHER INFORMATION: mim-1 protein
US-10-157-457-5

Query Match 9.2%; Score 74; DB 12; Length 22;
Best Local Similarity 61.9%; Pred. No. 0.042;
Matches 13; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 81 YONKAINNGVIRISGRFCVK 101
: | | | : | | | : | | | : | | |
Db 2 FHNQNAIDDDGVQISGSGYCVK 22

RESULT 14
US-09-884-570-6
Sequence 6, Application US/09884570
Patent No. US20020114779A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294
CURRENT APPLICATION NUMBER: US/09/884,570
CURRENT FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/212,271
PRIOR FILING DATE: 2000-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 6
LENGTH: 16
TYPE: PRT
ORGANISM: Gallus gallus
FEATURE:
LOCATION: 293..308
OTHER INFORMATION: amino acid sequence tryptic peptide from
OTHER INFORMATION: mim-1 protein
US-09-884-570-6

Query Match 9.1%; Score 73; DB 10; Length 16;
Best Local Similarity 73.3%; Pred. No. 0.037;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 129 VYPGIQSHVHENC 143
: | | | : | | | : | | | : | | |
Db 1 VFPGLIISHIVENC 15

RESULT 15
US-10-157-457-6
Sequence 6, Application US/10157457
Patent No. US20020150568A1
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294CIP
CURRENT APPLICATION NUMBER: US/10/157,457
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: US 09/884,570
PRIOR FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 6
LENGTH: 16
TYPE: PRT
ORGANISM: Gallus gallus
FEATURE:
LOCATION: 293..308
OTHER INFORMATION: amino acid sequence tryptic peptide from
OTHER INFORMATION: mim-1 protein
US-10-157-457-6

Query Match 9.1%; Score 73; DB 12; Length 16;
Best Local Similarity 73.3%; Pred. No. 0.037;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 129 VYPGIQSHVHENC 143
: | | | : | | | : | | | : | | |
Db 1 VFPGLIISHIVENC 15

Search completed: March 25, 2003, 17:29:44
Job time : 13 secs

C/Species: *Pichia angusta*
C/Date: 31-Mar-1993 #sequence_revision 30-Sep-1993 #text_change 20-Apr-2000
C/Accession: A23009; S30110
R:Janowicz, Z. A.; Eckart, M. R.; Drewe, C.; Roggenkamp, R. O.; Hollenberg, C. P.
Nucleic Acids Res. 13, 3043-3062, 1985
A/Title: Cloning and characterization of the DAS gene encoding the major methanol assim
A/Reference number: A23009; MUID:85215670; PMID:2987872
A/Accession: A23009
A/Molecule type: DNA
A/Residues: 1-693, 'RLRPEKGA' <JAN>
A/Cross-references: GB:X02424
A/Experimental source: ATCC 34438
A/Note: This sequence has been revised in reference S30110
R:Hansen, H.; Didion, T.; Thiemann, A.; Veenhuis, M.; Roggenkamp, R.
Mol. Gen. Genet. 235, 269-278, 1992
A/Title: Targeting sequences of the two major peroxisomal proteins in the methylotrophic
A/Reference number: S30110; MUID:93101130; PMID:1465101
A/Accession: S30110
A/Molecule type: DNA
A/Residues: 667-710 <HAN>
A/Cross-references: EMBL:X02424
A/Note: this is a revision to the sequence from reference A23009
C/Comment: This is the major methanol assimilatory enzyme from this methylotrophic organ
C/Genetics:
A/Gene: DAS
C/Superfamily: transketolase; thiamin pyrophosphate-binding domain homology
C/Keywords: peroxisome; thiamin pyrophosphate; transferase
F:18-208/Domian: thiamin pyrophosphate-binding domain homology #status atypical <TPB>
F:708-710/Region: peroxisome/glyoxysome location signal #status atypical

Query Match 9.1%; Score 73; DB 1; Length 710;
Best Local Similarity 24.1%; Pred. No. 18;
Matches 33; Conservative 21; Mismatches 49; Indels 34; Gaps 7;

Qy 11 GLISALAGPMANICGKSSNEIRTCORHCGQYSAORSPHOGVDXLCASGTYVAF 70
Db 525 GLFOKAVELPFSSITLS-LSNENLV-----QTLASRAQRNRNAGTILEAENAEVQI 575

Qy 71 TGMIVGOEKPYONKNAINNGVRISGRF-----CVKMFYIKPIKYKGP1--KKGEKL 120
Db 576 IG-VGAEMEFADKAK-----KILGRKFRTRVLSTPCTRLDEQSIGVRSVLRKMGGRV 628

Qy 121 GTLPLQKVPYPGIQSHV 137
Db 629 PTVV-----VDGHV 637

RESULT 11
T36565
probable penicillin-binding protein - Streptomyces coelicolor (fragment)
C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Nov-2000
C/Accession: T36565
R:Seeger, K. J.; Harris, D.; James, K. D.; Parkhill, J.; Barrell, B. G.; Rajandream, M. A.
submitted to the EMBL Data Library, June 1999
A/Reference number: Z21584
A/Accession: T36565
A/Status: preliminary; translated from GB/EMBL/DBDUT
A/Molecule type: DNA
A/Residues: 1-428 <SEP>
A/Cross-references: EMBL:AL079353; PTDN:CA84563.1; GSPDB:GN00070; SCOEDB:SCH17.14
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCOEDB:SCH17.14
C/Superfamily: penicillin-binding protein 1B

Query Match 9.0%; Score 72.5; DB 2; Length 428;
Best Local Similarity 38.9%; Pred. No. 12;
Matches 21; Conservative 6; Mismatches 22; Indels 5; Gaps 2;

Qy 43 QYASARSPHOGVDXLCASGTYVAFPTGMIV--GOEKPY--ONKNAINNGV 91
Db 342 QESVHQSLKDHVKNKSDVAASLIVPEGTGRITVAMGQSKPYGYGNTEINYSV 395

RESULT 12
AC2451
hypothetical protein all5163 [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp.
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C/Accession: AC2451
R:Kaneko, T.; Nakamura, Y.; Wolk, C. P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; I
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Ta
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacter
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AC2451
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-295 <KUR>
A/Cross-references: GB:BA000019; PTDN:BAB76862.1; PTD:917134301; GSPDB:GN00179
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: all5163

Query Match 8.9%; Score 71.5; DB 2; Length 295;
Best Local Similarity 29.6%; Pred. No. 11;
Matches 34; Conservative 14; Mismatches 40; Indels 27; Gaps 7;

Qy 53 HOGVDXLCASGTYV--YAPFTGMIVGOEKPYONKNAIN--NGVRISGRGFCVYKMFYIKPIK 109
Db 190 HSGVDLLALGTPVEALIPGTVEAFKQGYGNLVIIHNG-----GLQSRVYHLDSDIN 243

Qy 110 YK--GPVKKGEKIGTL---LPLQKVPYPGIQSHVHIENCSS-----DPIAYL 151
Db 244 KVGQKVNQGGOLIGTGTGQPTAK-----QPHLHFVRSVSSLCGVAENPKDYL 293

RESULT 13
T01351
subtilisin-like proteinase homolog F6N15.3 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 20-Sep-1999
C/Accession: T01351
R:Ryan, E.; Edwards, J.; Pape, K.
submitted to the EMBL Data Library, May 1998
A/Description: The sequence of A. thaliana F6N15.
A/Reference number: 214297
A/Accession: T01351
A/Status: translated from GB/EMBL/DBDUT
A/Molecule type: DNA
A/Residues: 1-706 <RYA>
A/Cross-references: EMBL:AF069299; MUID:93193311; PTD:93193320
A/Experimental source: cultivar Columbia
C/Genetics:
A/Map position: 4
A/Introns: 10/2; 43/1; 91/1; 126/2; 297/3; 330/2; 418/2; 454/3; 526/1
A/Note: F6N15.3
C/Superfamily: subtilisin-like proteinase ag12; subtilisin homology

Query Match 8.9%; Score 71.5; DB 2; Length 706;
Best Local Similarity 25.0%; Pred. No. 26;
Matches 33; Conservative 21; Mismatches 35; Indels 43; Gaps 8;

Qy 8 LLAG-----LISTALAGWANICAGKSSNEIRTCORHCGQYSAORSPHOGVD 57
Db 206 ILAGFEALIHGVEIISISIGPLADY-----SSDSI-----SVGSFHAMR-----KGIL 250

Qy 58 XLCASGTYVAFPTGMIV-----GOEKPYONKNAINNGVRISGRGFCVYKMFYI 105
Db 251 TVASAGND--GSSSGVTNHEPWLIVVAGSIDRFTFSKIDLGKNSFSGWG--ISWFSF 306

Qy 106 KPIKYKGPYKKG 117
Db 307 KAKSY--PLVSG 316

RESULT 14

E97836

probable periplasmic protein [imported] - Rickettsia conorii (strain Malish 7)

C:Species: Rickettsia conorii

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001

C:Accession: E97836

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R

Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A:Reference number: A97700; M01D:21442074; PMID:11557893

A:Accession: E97836

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-284 <KUR>

A:Cross-references: GB:AE006914; PIDN:AL03631.1; PID:915620216; GSPDB:GN00173

C:Genetics:

A:Gene: RC1093

Query Match

8.8%; Score 71; DB 2; Length 284;

Best Local Similarity 24.0%; Pred. No. 12;

Matches 24; Conservative 16; Mismatches 32; Indels 28; Gaps 5;

QY 53 HQVDXLCASGTVAPFTGMIV-GQEKPYQNKNAINGVIRISGRGFCVKMFTYIKPIRY 110

Db 181 HSGIDLOAKKAPITYAASGIVIKAAAPDYGNEVEIKH-----GRKFVTKYAHLEKMSV 235

QY 111 K-GPIKKGEKIGTLPLQKVPGIQ-----SHVHIE 140

Db 236 KEGNKIKRKQ-----FTIGQSTGNATGEHLHFE 264

RESULT 15

T05614

hypothetical protein F9D16.290 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 17-Mar-2000

C:Accession: T05614

R:Bevan, M.; Wedler, H.; Wedler, E.; Wambut, R.; Hohnselt, J.; Mewes, H.W.; Mayer, K.F.

submitted to the Protein Sequence Database, February 1999

A:Reference number: 215419

A:Accession: T05614

A:Molecule type: DNA

A:Residues: 1-444 <BEV>

A:Cross-references: EMBL:AL035394

A:Experimental source: cultivar Columbia; BAC clone F9D16

C:Genetics:

A:Map position: 4

A:Introns: 110/3; 151/1; 209/2; 233/1

A:Note: F9D16.290

C:Superfamily: polygalacturonase

Query Match

8.7%; Score 70; DB 2; Length 444;

Best Local Similarity 25.7%; Pred. No. 23;

Matches 38; Conservative 14; Mismatches 54; Indels 42; Gaps 8;

QY 23 NICAAGSSNEIRTC-----RHGCGOYSAORSOPHOGVDXLCASGTVAPFTGMIVGO 77

Db 237 NVCI--EDSYISTGDLVAIKSGMDQGIYVGRPSSNITIRITGS---SPFAGIATGS 290

QY 78 EKPYNKNAINGVIRISGRGFCVKM-----FYIKPIK---YKGP1KKGKELG----- 121

Db 291 ETSGIGIKNIIEHITLSNMGVGVNITKIRGIGIKNIKISDYVDYVDTAKYGIKIGADMGD 350

QY 122 -----TLPLQKVYPGIGQSHVHEN 141

Db 351 HPDENYNPALPVVK-----GIHITKN 371

Search completed: March 25, 2003, 17:22:09
Job time : 31.5 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:10:39 ; Search time 8.5 Seconds

(without alignments)
736.815 Million cell updates/sec

Title: US-09-868-953a-1

Perfect score: 804
Sequence: 1 MFSTKALLLAGLISTALAGP.....GIQSHVHENCDSDPAYL 151

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 802 | 99.8 | 151 | 1 | LCT2_HUMAN |
| 2 | 698 | 86.8 | 151 | 1 | LCT2_BOVIN |
| 3 | 681 | 84.7 | 151 | 1 | LCT2_MOUSE |
| 4 | 429 | 53.4 | 326 | 1 | MIM1_CHICK |
| 5 | 76.5 | 9.5 | 475 | 1 | YBBA_HAEN |
| 6 | 73 | 9.1 | 710 | 1 | DAS_PICAN |
| 7 | 71 | 8.8 | 779 | 1 | TRP3_THEVO |
| 8 | 70.5 | 8.8 | 680 | 1 | Z1M3_HUMAN |
| 9 | 69.5 | 8.6 | 472 | 1 | XDH_RAT |
| 10 | 69 | 8.6 | 1330 | 1 | Y763_CAEL |
| 11 | 68 | 8.5 | 184 | 1 | MUS2_AQUAE |
| 12 | 68 | 8.5 | 762 | 1 | COL1_ARATH |
| 13 | 67.5 | 8.4 | 535 | 1 | VCS0_HSVSA |
| 14 | 67.5 | 8.4 | 625 | 1 | TR1L_MOUSE |
| 15 | 67.5 | 8.3 | 398 | 1 | MOBI_XENLA |
| 16 | 67 | 8.1 | 1241 | 1 | NPHN_HUMAN |
| 17 | 65.5 | 8.1 | 271 | 1 | YBAF_ECOLI |
| 18 | 65 | 8.1 | 312 | 1 | PGLR_AGRTU |
| 19 | 65 | 8.1 | 317 | 1 | IBP2_SHEEP |
| 20 | 65 | 8.1 | 1388 | 1 | RPOD_TOBAC |
| 21 | 65 | 8.0 | 444 | 1 | FA7_RABIT |
| 22 | 64.5 | 8.0 | 571 | 1 | SECD_MYCLE |
| 23 | 64.5 | 8.0 | 2241 | 1 | TEGU_HCVNA |
| 24 | 64.5 | 8.0 | 593 | 1 | MTEP_HUMAN |
| 25 | 64 | 8.0 | 746 | 1 | GYP7_YEAST |
| 26 | 64 | 8.0 | 1477 | 1 | KEC2_DROME |
| 27 | 63.5 | 7.9 | 551 | 1 | LEM2_RABIT |
| 28 | 63 | 7.8 | 417 | 1 | KCRU_HUMAN |
| 29 | 63 | 7.8 | 490 | 1 | HOVD_CLOAM |
| 30 | 63 | 7.8 | 664 | 1 | PM5A_STRMU |
| 31 | 63 | 7.8 | 947 | 1 | PM19_CHLPN |
| 32 | 63 | 7.8 | 466 | 1 | FA7_HUMAN |
| 33 | 62.5 | 7.8 | | | |

| | | | | | | |
|----|------|-----|------|---|------------|---------------------|
| 34 | 62.5 | 7.8 | 691 | 1 | OAT6_HUMAN | O9y616 homo sapien |
| 35 | 62 | 7.7 | 109 | 1 | CYC_TERYX | P00079 tetranymena |
| 36 | 62 | 7.7 | 221 | 1 | GPI1_ZINOF | P82474 zingiber of |
| 37 | 62 | 7.7 | 274 | 1 | VPRP_MMTVB | P10271 mouse mamma |
| 38 | 62 | 7.7 | 499 | 1 | DHAS_CHICK | O93344 gallus galli |
| 39 | 62 | 7.7 | 655 | 1 | TR21_MOUSE | O9epus mus musculu |
| 40 | 62 | 7.7 | 775 | 1 | YTX1_XENLA | P14380 xenopus lae |
| 41 | 62 | 7.7 | 948 | 1 | HMDH_SCHMA | P16237 schistosoma |
| 42 | 62 | 7.7 | 1196 | 1 | BXCN_CLOBO | P46081 clostridium |
| 43 | 62 | 7.7 | 2109 | 1 | RRLP_VSVUH | P13615 vesicular s |
| 44 | 62 | 7.7 | 3412 | 1 | POUG_TREVS | P07720 t genome po |
| 45 | 61.5 | 7.6 | 362 | 1 | FIBP_ADE08 | P36845 human adeno |

ALIGNMENTS

RESULT 1
LCT2_HUMAN STANDARD; PRT; 151 AA.
AC O14960; O14565;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Leukocyte cell-derived chemotaxin 2 precursor (hLECT2).
GN LECT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT ILE-58.
RC TISSUE=Liver;
RC MEDLINE=98193133; PubMed=9524238;
RA Yamagoe S., Mizuno S., Suzuki K.;
RT "Molecular cloning of human and bovine LECT2 having a neutrophil
RT chemotactic activity and its specific expression in the liver.";
RL Blochim. Biophys. Acta 1396:105-113(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripherel blood;
RC MEDLINE=98207247; PubMed=9545637;
RA Yamagoe S., Kameoka Y., Hashimoto K., Mizuno S., Suzuki K.;
RT "Molecular cloning, structural characterization, and chromosomal
RT mapping of the human LECT2 gene.";
RL Genomics 48:324-329(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=99048897; PubMed=9832057;
RA Nagai H., Hamada T., Uchida T., Yamagoe S., Suzuki K.;
RT "systemic expression of a newly recognized protein, LECT2, in the
RT human body.";
RL Pathol. Int. 48:882-886(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC Kimmerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,
RA Kadner K., Miguel T., Miller C., Pittluck S., Pollard M., Rojeski H.,
RA Subramanian S., Martin C.H.;
RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC Strong C., Biewald T., Tin-Wollam A., Duckels G.;
RA Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SUBCELLULAR LOCATION.
RC MEDLINE=97031488; PubMed=8877413;
RA Yamagoe S., Yamakawa Y., Matsuo Y., Minowada J., Mizuno S., Suzuki K.;
RT "Purification and primary amino acid sequence of a novel neutrophil
RT chemotactic factor LECT2.";
CC Immunol. Lett. 52:9-13(1996).
CC -!- FUNCTION: Has a neutrophil chemotactic activity. Also a positive
CC regulator of chondrocyte proliferation.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and secreted.

CC -1- TISSUE SPECIFICITY: Highly expressed in adult and fetal liver and
 CC weakly in testis. Not expressed in bone marrow.
 CC -1- INDUCTION: By phytohemagglutinin (PHA).
 CC -1- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D63521; BAA23609.1; -;
 DR EMBL: AB007546; BAA25669.1; -;
 DR EMBL: AC004763; AAC17734.1; -;
 DR EMBL: AC002428; AAB66905.1; -;
 DR Genew; HGNC:6550; LECT2.
 DR MIM; 602882; -;
 DR Chemotaxis; Signal; Polymorphism.
 KW SIGNAL 1 18 BY SIMILARITY.
 FT CHAIN 19 151 LEUCOCYTE CELL-DERIVED CHEMOTAXIN 2.
 FT VARIANT 58 58 V->I.
 SQ SEQUENCE 151 AA; 16376 MW; A4267F0A24B69631 CRC64;
 /FTD-VAR-011386;
 Query Match 99.8%; Score 802; DB 1; Length 151;
 Best Local Similarity 99.3%; Pred. No. 3, 1e-77;
 Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MESTKALLAGLITSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAORSQRPBGVDXLC 60
 DB 1 MESTKALLAGLITSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAORSQRPBGVDXLC 60
 QY 61 SAGSTVYAPFTGMIVGQEKPYONKNAIINGVRISGRGFCVKKFYIKPIKPKYKGGKRL 120
 DB 61 SAGSTVYAPFTGMIVGQEKPYONKNAIINGVRISGRGFCVKKFYIKPIKPKYKGGKRL 120
 QY 121 GTLLPQKVPYPGIOSHVHIENCDSPTAYL 151
 DB 121 GTLLPQKVPYPGIOSHVHIENCDSPTAYL 151
 RESULT 2
 LCT2_BOVIN STANDARD; PRT; 151 AA.
 ID LCT2_BOVIN
 AC O62644;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leukocyte cell-derived chemotaxin 2 precursor (bLect2) (Chondromodulin
 DE II) (bChm-II).
 GN LECT2.
 OS Bos taurus (Bovine).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 OX NCBI_TaxID=9913;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=98193133; PubMed=9524238;
 RA Yamagoe S., Mizuno S., Suzuki K.;
 RT "Molecular cloning of human and bovine LECT2 having a neutrophil
 RT chemotactic activity and its specific expression in the liver.";
 RL Blochim. Biophys. Acta 1396:105-113(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=99160594; PubMed=10050029;
 RA Shukunami C., Kondo J., Takahashi K., Inoue H., Kamizono A.,
 RA Hiraki Y.;
 RT "Molecular cloning of mouse and bovine chondromodulin-II cDNAs and the

RT growth-promoting actions of bovine recombinant protein.";
 RL J. Biochem. 125:436-442(1999).
 RN [3]
 RP SEQUENCE OF 19-151.
 RC TISSUE=Epiphyseal cartilage;
 RX MEDLINE=96394331; PubMed=8798437;
 RA Hiraki Y., Inoue H., Kondo J., Kamizono A., Yoshitake Y.,
 RA Shukunami C., Suzuki F.;
 RT "A novel growth-promoting factor derived from fetal bovine cartilage,
 RT chondromodulin II. Purification and amino acid sequence.";
 RL J. Biol. Chem. 271:22657-22662(1996).
 CC -1- FUNCTION: Has a neutrophil chemotactic activity. Also a positive
 CC regulator of chondrocyte proliferation.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB001350; BAA25302.1; -;
 DR EMBL: D89011; BAB18616.1; -;
 KW Chemotaxis; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 151 LEUCOCYTE CELL-DERIVED CHEMOTAXIN 2.
 FT VARIANT 19 151
 SQ SEQUENCE 151 AA; 16320 MW; 2A1D8F5B28A3B8 CRC64;
 Query Match 86.8%; Score 698; DB 1; Length 151;
 Best Local Similarity 86.1%; Pred. No. 2, 6e-66;
 Matches 130; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
 QY 1 MESTKALLAGLITSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAORSQRPBGVDXLC 60
 DB 1 MESTKALLAGLITSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAORSQRPBGVDXLC 60
 QY 61 SAGSTVYAPFTGMIVGQEKPYONKNAIINGVRISGRGFCVKKFYIKPIKPKYKGGKRL 120
 DB 61 SAGSTVYAPFTGMIVGQEKPYONKNAIINGVRISGRGFCVKKFYIKPIKPKYKGGKRL 120
 QY 121 GTLLPQKVPYPGIOSHVHIENCDSPTAYL 151
 DB 121 GTLLPQKVPYPGIOSHVHIENCDSPTAYL 151
 RESULT 3
 LCT2_MOUSE STANDARD; PRT; 151 AA.
 ID LCT2_MOUSE
 AC O88803; O88804; O9QWN3; Q9Z337;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leukocyte cell-derived chemotaxin 2 precursor (Chondromodulin II)
 DE (Chm-II).
 GN LECT2.
 OS Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 OX NCBI_TaxID=10090;
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANT VAL-129.
 RC STRAIN=BALB/C; TISSUE=Liver;
 RX MEDLINE=98382586; PubMed=9714793;
 RA Yamagoe S., Watanabe T., Mizuno S., Suzuki K.;
 RT "The mouse Lect2 gene: cloning of cDNA and genomic DNA, structural
 RT characterization and chromosomal localization.";
 RL Gene 216:171-178(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=Swiss Webster / NIH Swiss; TISSUE=Embryo, and Liver;

```

RX MEDLINE=99160594; PubMed=10050029;
RA Shukunami C., Kondo J., Wakai H., Takahashi K., Inoue H., Kamizono A.,
RA Hiraki Y.;
RT "Molecular cloning of mouse and bovine chondromodulin-II cDNAs and the
RT growth-promoting actions of bovine recombinant protein.";
RL J. Biochem. 125:436-442(1999)
CC
CC -1- FUNCTION: Has a neutrophil chemotactic activity. Also a positive
CC regulator of chondrocyte proliferation.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1/LECT2 (shown here) and
CC 2/LECT20; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Highly expressed in liver and weakly in
CC testis. Not expressed in heart, brain, spleen, lung, skeletal
CC muscle and kidney.
CC -1- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AB009687; BAA33383.1; -
DR EMBL: AB009688; BAA33384.1; -
DR EMBL: AB009689; BAA33385.1; -
DR EMBL: AB009689; BAA33386.1; -
DR EMBL: AF035161; AF13302.1; -
DR MGD: MG1:1278342; Lect2.
KW Chemotaxis; Signal; Alternative splicing.
FT SIGNAL 1
FT CHAIN 19 151
FT VARSPLIC 98 151
FT
FT VARIANT 129 129
FT SEQUENCE 151 AA; 16405 MW; 18AF444046B7AE8E CRC64;
SQ
Query Match 84.7%; Score 681; DB 1; Length 151;
Best Local Similarity 82.1%; Pred. No. 1.6e-64;
Matches 124; Conservative 12; Mismatches 15; Indels 0; Gaps 0;
QY 1 MESTKALLAGLSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAGRSORPHOGVDXLC 60
DB 1 MPTTLLISALLSALLSALSPWANICAGKSSNEIRTCDSYCGCGYSAGRTGRHHPGDVILC 60
QY 61 SAGSTVAPFTGMVGOEKPYONKNAINGVRIISGRGFCYKMYIKRIRKPKIKKPKK 120
DB 61 SDGSVYAPFTGKTIVGOKPYRNKNAINDGIRLSGRGFCVYIKRIRKPKIKKSGIKK 120
QY 121 GTLLPLOKVVYQIASHVHIENCDSDPTAYL 151
DB 121 GTLLPLOKITYPGIASHVHVENCDSDPTAYL 151

```

RESULT 4
MIM1_CHICK
ID MIM1_CHICK STANDARD; PRT: 326 AA.

AC P08940;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Myeloid protein-1 precursor (P33).
GN MIM-1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90090611; PubMed=2688896;

```

RA Ness S.A., Marknell A., Graf T.;
RT "The v-myb oncogene product binds to and activates the promyelocyte-
RT specific mim-1 gene.";
RL Cell 59:1115-1125(1989).
CC
CC -1- FUNCTION: NOT KNOWN.
CC -1- SUBCELLULAR LOCATION: GRANULES OF PROMYELOCYTES.
CC -1- INDUCTION: BY THE MYB ONCOGENE.
CC -1- PM: SUBSTRATE FOR ARGININE-SPECIFIC ADP-RIBOSYLTRANSFERASE.
CC -1- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M29448; AAA48954.1; -
DR EMBL: M29449; AAA48958.1; -
DR PIR: A33755; A33755.
DR PIR: S29229; S29229.
KW Repeat; Signal.
FT SIGNAL 1 23
FT CHAIN 24 326
FT REPEAT 28 162
FT REPEAT 177 312
FT REPEAT 297 297
FT CONFLICT 297 297
FT SEQUENCE 326 AA; 35636 MW; B8C4742EF0D3BAD3 CRC64;
SQ
Query Match 53.4%; Score 429; DB 1; Length 326;
Best Local Similarity 50.3%; Pred. No. 1.1e-37;
Matches 80; Conservative 26; Mismatches 39; Indels 14; Gaps 2;
QY 7 LLAGLSTALAGP-----WANICAGKSSNEIRTCDRHGGCGQYSAGRSOR--P 52
DB 7 IALLSVSTARARQGEVHPDQGGRRMAQISGNPFRIRIGCDRYGCGNYCASRGKKEK 66
QY 53 HOGVDXLCASAGSTVAPFTGMVGOEKPYONKNAINGVRIISGRGFCYKMYIKRIRKPK 112
DB 67 HKGVAVICTDGSIVYAPPSGQLSGPIRFFHNGAIDGQVQISGSGCYKLVCIHPIRYHG 126
QY 113 PIRKGEKLTLLPLOKVVYQIASHVHIENCDSDPTAYL 151
DB 127 QIQKGQOLGRMLPMQKVEPGIYSHVHVENCDSDPTLL 165

```

RESULT 5
YERBA_HAEIN
ID YERBA_HAEIN STANDARD; PRT: 475 AA.

AC P44653;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein HI0409.
GN HI0409.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;
OC Haemophilus.
OX NCBI_TaxID=727;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / K120 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodok A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
 RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RL Rd.";
 RL Science 269:496-512(1995).
 CC -1- FUNCTION: COULD BE INVOLVED IN CELL WALL DEGRADATION OR
 CC FORMATION (BY SIMILARITY).
 CC -1- SIMILARITY: STRONG, TO E.COLI YEBA.
 CC -1- SIMILARITY: TO STAPHYLOCOCCUS LYSOSTAPHIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL: U32724; AAC22068.1; -.
 DR MEROPS: M37.UFW; -.
 DR TIGR: H10409; -.
 DR InterPro: IPR002886; Peptidase_M37.
 DR Pfam: PF01551; Peptidase_M37.1.
 DR Hypothetical protein: Cell wall; Complete proteome.
 KW Hypothetical protein: Cell wall; 143C10F9233939D CRC64;
 SQ SEQUENCE 475 AA: 53255 MW: 143C10F9233939D CRC64;
 Query Match 9.5%; Score 76.5; DB 1; Length 475;
 Best Local Similarity 28.7%; Pred. No. 1.7;
 Matches 27; Conservative 15; Mismatches 41; Indels 11; Gaps 4;
 QY 51 RPHGVADMLCSAGSTYAPFTGMIVGOEKPYONKNAINNGVRISGRGCVKMFYIKP-I 108
 DB 346 RPHKGVADSVSOGTPVIAADGTIV--EKVAVYQAGAGRYVLMRGREQVTVYMLSKSLV 403
 QY 109 KYKPIKKGKGLTLLPLQKVYPIGOS--HVHIE 140
 DB 404 KAGOTVKKGERIAL-----SGNTGISTGPHDHYE 432
 RESULT 6
 DAS_PICAN STANDARD; PRT; 710 AA.
 ID DAS_PICAN
 AC P06834;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Dihydroxy-acetone synthase (EC 2.2.1.3) (DHAS) (formaldehyde
 DE transketolase) (glycerone synthase).
 GN DAS.
 OS Pichia angusta (Yeast) (Hansenula polymorpha).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4905;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 34438;
 RX MEDLINE=85215670; PubMed=2987872;
 RA JENOWICZ Z.A., Eckart M.R., Drewke C., Roggenkamp R.O.,
 RA Hollenberg C.P., Maat J., Ledebor A.M., Visser C., Verrips C.T.;
 RT "Cloning and characterization of the DAS gene encoding the major
 RT methanol assimilatory enzyme from the methylotrophic yeast Hansenula
 RT polymorpha.";
 RL Nucleic Acids Res. 13:3043-3062(1985).
 RN [2]
 RP REVISIONS, SEQUENCE OF 667-710 FROM N.A.
 RC STRAIN-ATCC 34438;

RX MEDLINE=93101130; PubMed=1465101;
 RA Hansen H., Didion T., Thiemann A., Veenhuis M., Roggenkamp R.O.;
 RT "Targeting sequences of the two major peroxisomal proteins in the
 RT methylotrophic yeast Hansenula polymorpha.";
 RL Mol. Gen. Genet. 235:269-278(1992).
 CC -1- FUNCTION: THIS IS THE MAJOR METHANOL ASSIMILATORY ENZYME FROM
 CC THE METHYLOTROPHIC HANSENULA POLYMORPHA.
 CC -1- CATALYTIC ACTIVITY: D-xylulose 5-phosphate + formaldehyde =
 CC glyceraldehyde 3-phosphate + glycercione.
 CC -1- COFACTOR: THIAMINE PYROPHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: Peroxisomal.
 CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL: X02424; CAA26276.1; -.
 DR PIR: A23009; XJHOFK.
 DR HSP: P23254; 1TRK.
 DR InterPro: IPR000360; Transketolase.
 DR Pfam: PF00456; transketolase; 1.
 DR Pfam: PF02779; transket-pyr; 1.
 DR Pfam: PF02780; transketolase_C; 1.
 DR PROSITE: PS00801; TRANSKETOLASE_1; 1.
 DR PROSITE: PS00802; TRANSKETOLASE_2; 1.
 DR PROSITE: PS00342; MICROBODIES_CTER; 1.
 DR Methanol utilization; Peroxisome; transferase; Thiamine pyrophosphate.
 KW SITE 708 710 MICROBODIES TARGETING SIGNAL (POTENTIAL).
 FT SITE 708 710
 SQ SEQUENCE 710 AA: 78842 MW: E424B6D37CC8B0B CRC64;
 Query Match 9.1%; Score 73; DB 1; Length 710;
 Best Local Similarity 24.1%; Pred. No. 6;
 Matches 33; Conservative 21; Mismatches 49; Indels 34; Gaps 7;
 QY 11 GLISTALGPMNINICAGKSSNEIRTCDRHCGOYSAQRSPRHGVADMLCSAGSTYAPF 70
 DB 525 GLFQKAVELPFSSILS-LSRNEVL-----GLASRAQRRNAGYILEDAENAEVQI 575
 QY 71 TGMIVGOEKPYONKNAINNGVRISGRG-----CYKMFYPIKIKKPI--KKGEVL 120
 DB 576 IG--VGAEMEFADKA-----KILGRKFRTRVLSIPETRLPDEDSIGYRSVLRDGRGV 628
 QY 121 GTLLPLQKVYPIGOSHV 137
 DB 629 PTVV-----VDGHV 637
 RESULT 7
 TRF3_THEVO STANDARD; PRT; 779 AA.
 ID TRF3_THEVO
 AC Q97AJ6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tricorn protease interacting factor F3 (EC 3.4.11.-).
 GN TRF3 OR TV0814 OR TVG0817891.
 OS Thermoplasma volcanium.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 OC Thermoplasmataceae; Thermoplasma.
 OX NCBI_TaxID=50339;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GSSI / DSM 4299 / JCM 9571;
 RX MEDLINE=20570466; PubMed=11121031;
 RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
 RA Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
 RA Nunoshita T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
 RT "Archaeal adaptation to higher temperatures revealed by genomic

RT sequence of Thermoplasma volcanium.*;
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
 CC -1- FUNCTION: Proteases F1, F2 and F3 degrade oligopeptides produced
 CC by tricorn (themselves probably produced by the proteasome),
 CC yielding free amino acids (By similarity).
 CC -1- COFACTOR: Binds one zinc ion (By similarity).
 CC -1- SUBUNIT: Part of the tricorn proteolytic complex (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AP000993; BAB59956.1; -.
 DR MEROPS: M01.021; -.
 DR InterPro: IPR001930; Ala-peptase.
 DR InterPro: IPR000130; Zn.Mpeptase.
 DR Pfam: PF01433; Peptidase.M1.1.
 DR PRINTS: PR00756; ALADIPRASE.
 DR PROSITE: PS00142; ZINC_PROTEASE.1.
 DR HYDROLASE: Metalloprotease; Aminopeptidase; Zinc; Complete proteome.
 KW METAL 266 266 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 267 267 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 270 270 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 289 289 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 352 352 PROTON DONOR (POTENTIAL).
 SQ SEQUENCE 779 AA; 88068 MW; 6BC227FE2ECCECB201 CRC64;
 Query Match 8.8%; Score 71; DB 1; Length 779;
 Best local similarity 26.4%; Pred. No. 11;
 Matches 33; Conservative 21; Mismatches 59; Indels 12; Gaps 4;
 QY 24 ICAGKSSNEIRTCORHGGQGYAQRSGRPHQGVXLCASGTYVAPFTGMVGGEKPYQN 83
 DB 590 IAYGKASRLALLDEAYC-ETLAPRFSNFEQCPPELKSATATVALSTGDKVMEKYS 648
 QY 84 KNAINNCVR-ISGRG---FCVGMFYIKPIKYGPIKKEKLG-----TLLPQKVP 132
 DB 649 LDRDEDYKIIISGFELKSSIDLSVYSGMIERGEIKQDMLSEYLSALETWAGREYISN 708
 QY 133 IQSHV 137
 DB 709 LENTV 713
 RESULT 8
 2334_HUMAN STANDARD; PRT; 680 AA.
 ID 2334_HUMAN STANDARD; PRT; 680 AA.
 AC 09HCL1; 09NYM4;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Zinc finger protein 334.
 GN ZNF334.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=1638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grahm D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Levasialho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurtry A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mollikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel A., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prithalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Stuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swan R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmshurst L., Wray P.W., Hubbard T., Dudin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.*";
 RL Nature 414:865-871(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Nagatsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
 RA Nimomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project.*";
 RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Strausberg R.;
 RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -----
 CC -1- SIMILARITY: CONTAINS 1 KRAB DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AL034424; CAC18856.1; -.
 DR EMBL: AL354745; CAC03544.1; -.
 DR EMBL: AK001331; BA091630.1; -.
 DR EMBL: BC024177; AAH24177.1; -.
 DR HSSP: P03001; ITP6.
 DR Genew: HGNC:15806; ZNF334.
 DR InterPro: IPR001909; KRAB.
 DR InterPro: IPR000822; Znf.C2H2.
 DR Pfam: PF000096; 2f.C2H2; 28.
 DR Pfam: PF01352; KRAB; 2.
 DR ProDom: PD000003; Znf.C2H2; 7.
 DR SMART: SM00349; KRAB; 1.
 DR SMART: SM00355; Znf.C2H2; 14.
 DR PROSITE: PS50805; KRAB; 1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 14.
 DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 14.
 KW Transcription regulation; Zinc-finger; Metal-binding; Nuclear protein;
 KW DNA-binding; Repeat.
 FT DOMAIN 10 81 KRAB.
 FT ZINC_FINGERS.
 FT ZN_FING 237 259 C2H2-TYPE.
 FT ZN_FING 265 287 C2H2-TYPE.
 FT ZN_FING 293 315 C2H2-TYPE.

FT ZN_FING 321 343 C2H2-TYPE.
 FT ZN_FING 349 371 C2H2-TYPE.
 FT ZN_FING 377 399 C2H2-TYPE.
 FT ZN_FING 405 427 C2H2-TYPE.
 FT ZN_FING 433 455 C2H2-TYPE.
 FT ZN_FING 461 483 C2H2-TYPE.
 FT ZN_FING 544 566 C2H2-TYPE.
 FT ZN_FING 572 594 C2H2-TYPE.
 FT ZN_FING 600 622 C2H2-TYPE.
 FT ZN_FING 628 650 C2H2-TYPE.
 FT ZN_FING 656 678 C2H2-TYPE.
 FT CONFLICT 118 118 K -> E (IN REF. 2).
 FT CONFLICT 426 426 S -> G (IN REF. 2).
 SQ SEQUENCE 680 AA: 79648 MW: 77DB5086B9261DC6 CRC64;

Query Match 8.8%; Score 70.5; DB 1; Length 680;
 Best Local Similarity 21.4%; Pred. No. 11;
 Matches 28; Conservative 15; Mismatches 45; Indels 43; Gaps 4;

27 GKSNIEIRTCDRHGGCGOYSAORSQRPHQVDLCSAGSTVYAPFTGMIVGQEPYONKNA 86
 568 GORPECNCGKTFCCQKFSVEHQRTHTG-----EKPYECNCG 605
 87 INNGVIRISGRGFCVKMPY-----IKPIKYGPRIKGEKLTLLPLQKYVPGIQSHV 137
 606 -----GKSFCHKSAFVYHRRIRHTGEKPEYCNQCGKTYRRLMTLTHQKIHTEGKPY- 656

QY 138 HIENCSSDPT 148
 Db 657 ---ECKCKERT 664

RESULT 9
 ZIM3_HUMAN
 ID ZIM3_HUMAN STANDARD; PRT; 472 AA.
 AC Q96PE6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Zinc finger imprinted 3.
 GN ZIM3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21429118; PubMed=11543637;
 RA Kim J., Bergmann A., Wehrli E., Lu X., Stubbs L.,
 RT "Imprinting and evolution of two Kruppel-type zinc-finger genes, ZIM3
 and ZNF264, located in the PEG3/USP29 imprinted domain.";
 RL Genomics 77:91-98(2001).
 CC -1- FUNCTION: May function as a transcription factor.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 FINGER PROTEINS.
 CC -1- SIMILARITY: CONTAINS 1 KRAB DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: AF365931; AAL11635.1; -
 DR Genbank: HGNC:16366; ZIM3.
 DR InterPro: IPR001909; KRAB.
 DR InterPro: IPR000822; ZnfC2H2.
 DR Pfam: PF00096; zf-C2H2; 1.
 DR Pfam: PF01352; KRAB; 1.
 DR PROSITE: PS50805; KRAB; 1.

DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 11.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 11.
 KW Transcription regulation; Zinc-finger; Metal-binding; Nuclear protein;
 KW DNA-binding; Repeat.
 FT DOMAIN 8 80 KRAB.
 FT ZN_FING 167 470 ZINC_FINGERS.
 FT ZN_FING 167 189 C2H2-TYPE.
 FT ZN_FING 195 217 C2H2-TYPE.
 FT ZN_FING 223 245 C2H2-TYPE.
 FT ZN_FING 251 273 C2H2-TYPE.
 FT ZN_FING 279 301 C2H2-TYPE.
 FT ZN_FING 307 329 C2H2-TYPE.
 FT ZN_FING 335 357 C2H2-TYPE.
 FT ZN_FING 363 385 C2H2-TYPE.
 FT ZN_FING 391 413 C2H2-TYPE.
 FT ZN_FING 419 441 C2H2-TYPE.
 FT ZN_FING 447 470 C2H2-TYPE.
 SQ SEQUENCE 472 AA: 54498 MW: 755279B78653F286 CRC64;

Query Match 8.6%; Score 69.5; DB 1; Length 472;
 Best Local Similarity 24.1%; Pred. No. 9.1;
 Matches 34; Conservative 19; Mismatches 67; Indels 21; Gaps 5;

1 MSTKALLAGLSTALAGWANTCARKSSNIEITCDRHGGCGOYSAORSQRPHQVDLXIC 60
 175 LFSKSRSLQSHLRHACQKPECHSGRAEGEKWLDRH---QKTHAEKPYK-----C 225
 61 SAGSTVYAPFTGMIVGCO-----EKPYONKNAINGVIRISGRGFCVKMPYIKPIKYGPRIK 115
 226 EMCGNAYKOKSNLFOHOKMTKEKPYCKTC---GKAFSWSSCINHEKHNAKSYQCN 282
 116 KGEKL---GTLPLQKYVPG 132
 283 ECKSFQNSTLQHKKVVHG 303

QY 116 KGEKL---GTLPLQKYVPG 132
 Db 283 ECKSFQNSTLQHKKVVHG 303

RESULT 10
 XDH_RAT
 ID XDH_RAT STANDARD; PRT; 1330 AA.
 AC P22985; O63157;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Xanthine dehydrogenase/oxidase (includes: Xanthine dehydrogenase
 (EC 1.1.1.204) (XD); Xanthine oxidase (EC 1.1.3.22) (XO) (Xanthine
 oxidoreductase)].
 DE XDH.
 GN Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE: Liver.
 RX MEDLINE=90354396; PubMed=2387845;
 RA Amaya Y., Yamazaki K.-I., Sato M., Noda K., Nishino T., Nishino T.;
 RT "Proteolytic conversion of xanthine dehydrogenase from the
 RT NAD-dependent type to the O2-dependent type. Amino acid sequence of
 RT rat liver xanthine dehydrogenase and identification of the cleavage
 RT sites of the enzyme protein during irreversible conversion by
 RT trypsin.";
 RL J. Biol. Chem. 265:14170-14175(1990).
 RN [2]
 RP SEQUENCE OF 1-54 FROM N.A.
 RC STRAIN: Sprague-Dawley;
 RX MEDLINE=94268906; PubMed=8208609;
 RA Chow C.W., Clark M., Rinaldo J., Chalkley R.;
 RT "Identification of the rat xanthine dehydrogenase/oxidase promoter.";
 RL Nucleic Acids Res. 22:1846-1854(1994).
 CC -1- FUNCTION: THIS ENZYME CAN BE CONVERTED FROM THE DEHYDROGENASE FORM
 CC (D) TO THE OXIDASE FORM (O) IRREVERSIBLY BY PROTEOLYSIS OR
 CC REVERSIBLY THROUGH THE OXIDATION OF SULFHYDRYL GROUPS.
 CC -1- CATALYTIC ACTIVITY: Xanthine + NAD(+) + H(2)O = urate + NADH.

```

CC -1- CATALYTIC ACTIVITY: Xanthine + H(2)O + O(2) = urate + H(2)O(2).
CC -1- COFACTOR: FAD, MOLYBDOTERIN, AND TWO 2FE-2S CLUSTERS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Peroxisomal.
CC -1- INDUCTION: BY INTERFERON.
CC -1- SIMILARITY: BELONGS TO THE XANTHINE DEHYDROGENASE FAMILY.
CC -1- SIMILARITY: TO 2FE-2S FERRODOXINS IN THE N-TERMINAL DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: J05579; AAA42349.1; -.
DR EMBL: U08122; AAA18869.1; -.
DR EMBL: U08120; AAA18869.1; JOINED.
DR EMBL: U08121; AAA18869.1; JOINED.
DR PIR: A37810; A37810.
DR HSSP: P80457; 1FIO.
DR InterPro: IPR002888; 2Fe-2S_bind.
DR InterPro: IPR000564; 2Fe2S_ferredoxin.
DR InterPro: IPR000674; Aldxan_dh.C.
DR InterPro: IPR0005107; CO_deh_flav.C.
DR InterPro: IPR0005072; Euk_MB_oxred.
DR InterPro: IPR001041; Ferredoxin.
DR InterPro: IPR002346; dehydrog_molyb.
DR Pfam: PF00111; fer2.1.
DR Pfam: PF00941; FAD_binding-5; 1.
DR Pfam: PF01315; Ald_xan_dh.C; 1.
DR Pfam: PF01799; fer2.2; 1.
DR Pfam: PF02738; Ald_xan_dh.C2; 1.
DR Pfam: PF03450; CO_deh_flav.C; 1.
DR ProDom: PD16071; 2Fe-2S_bind.1.
DR ProSite: PS00197; 2FE2S_FERREDOXIN.1.
DR ProSite: PS00559; MOLYBDOTERIN_EUK; 1.
DR Oxidoreductase; NAD; Molybdenum; Flavoprotein; FAD; Iron-sulfur.
FT INIT_MET 0 0
FT METAL 36 36 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 42 42 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 47 47 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 50 50 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
SQ SEQUENCE 1330 AA; 146111 MW; A3DD206B9D74E565 CRC64;

Query Match 8.6%; Score 69; DB 1; Length 1330;
Best Local Similarity 28.0%; Pred. No. 31;
Matches 33; Conservative 12; Mismatches 43; Indels 30; Gaps 6;

QY 35 TCDRHGCGGYSA---QRSGRPHQGVDXLCSAGSTVYAPFTGMIVGOEKPYNKNKINNGV 91
| | | | | : | | | | | : | | | | |
Db 660 TCYGHIIIGAVVADPEHQRARGV-----KITTEDLPAILTIQD-----AINN-- 703

QY 92 RISGRGCVKMYIKPIKY-KGPIKKGKELGTLPLQKVVPGIOSHVAIE-NCDSSDP 147
| | | | | : | | | | | : | | | | |
Db 704 -----NSFYGSEIKIEKGLDKKGFSEADNVVSGELYIGQEHRYLEENCTIAPV 752

RESULT 11
Y763.CAEEL STANDARD; PRT; 184 AA.
AC 011079;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 20.9 kDa protein B0563.3 in chromosome X.
GN B0563.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA STRAIN-Bristol N2;
RA Favella T.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE B11 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U28740; AAA68321.1; -.
DR WormPep; B0563.3; CE02441.
DR InterPro: IPR002199; Bax_inhbrl.
DR Pfam: PF01027; UPF0005; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 70 90 POTENTIAL.
FT TRANSMEM 96 116 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
FT TRANSMEM 160 180 POTENTIAL.
SQ SEQUENCE 184 AA; 20942 MW; 09836D218202F378 CRC64;

Query Match 8.5%; Score 68; DB 1; Length 184;
Best Local Similarity 27.2%; Pred. No. 4.8;
Matches 31; Conservative 9; Mismatches 32; Indels 42; Gaps 7;

QY 25 CAGKSSNEIRTCRHGCGGYSAO--RSQRPHQGVDXLCSAGS-----TVY--- 67
| | | | | : | | | | | : | | | | |
Db 21 CVPKSS-----RFGSCDYCAPRLRQSRSSQ---LCSGRHRCSSCHYGMILTFEAK 69

QY 68 -----APPTGMIVGOEKP--QKNKALNNKVRISGRFCV-----KMFYIKP 107
| | | | | : | | | | | : | | | | |
Db 70 VLEAAVITGLVVASLFAVTLQNKRDSEVSGASGLICVLLMAGIFQMFPMSP 123

RESULT 12
MUS2.AOUAE STANDARD; PRT; 762 AA.
AC 067287;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MUS2 protein.
DE MUS2B OR MUS22 OR AQ_1242.
GN Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
OC Aquifex.
OX NCBI_TaxID=63363;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
RL -1- FUNCTION: NOT KNOWN.
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC

```

DR EMBL: AE000730; AAC07247.1; -.
 DR InterPro: IPR000302; KOW_motif.
 DR InterPro: IPR000432; Muts_C.
 DR InterPro: IPR002663; Muts_N.
 DR InterPro: IPR002625; Smr/Muts2.
 DR Pfam: PF00467; KOW; 1.
 DR Pfam: PF00488; Muts_C; 1.
 DR Pfam: PF01713; Smr; 1.
 DR ProDom: PD001263; Muts_C; 1.
 DR SMART: SM00533; MutsC; 1.
 DR SMART: SM00533; MutsC; 1.
 DR SMART: SM00463; Smr; 1.
 DR TIGRFAMs: TIGR01069; muts2; 1.
 DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; FALSE_NEG.
 KW ATP-binding; DNA-binding; Complete proteome.
 FT NP_BIND 329 336 ATP (POTENTIAL).
 SQ SEQUENCE 762 AA; 87538 MW; 3A7D977DC2EB01BC CRC64;

Query Match 8.5%; Score 68; DB 1; Length 762;
 Best Local Similarity 21.8%; Pred. No. 22;
 Matches 27; Conservative 19; Mismatches 40; Indels 38; Gaps 4;

QY 64 STVYAFETGIVGQEKRYOKNAINNGVRISGRF---CYKMFYIKPIKYKPIKGEK 119
 DB 20 SYAHSPATEKTIQNLKPYTKRKYKEIEISKAFEDIANVRLFFEDIREL--LKKAKL 77
 QY 120 LGTL-----PLQVYEGIOSHVHIEN--CDSSDP 147
 DB 78 QGAILGVEDIKILNINILTKETIRVLSHVORLEPLKRYKKLYTFPLENLIISIDP 137
 QY 148 TAYL 151
 DB 138 RGFV 141

RESULT 13

COLI_ARATH STANDARD; PRT; 355 AA.
 AC 050055;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Zinc finger protein CONSTANS-LIKE 1.
 GN COLI OR A55G15850 OR F14F8_230.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Eucotids II; Brassicales; Brassicaceae; Arabidopsis.
 CX NCBI_TaxID=3702;

SEQUENCE FROM N.A.
 RP STRAIN=cv. Landsberg erecta; TISSUE=Aerial parts;
 RA Puterill J.J., Ledger S.E., Lee K., Robson F., Murphy G.,
 RA Coupland G.;
 RT "The flowering-time gene CONSTANS and homologue CONSTANS LIKE 1 exist
 as a tandem repeat on chromosome 5 of Arabidopsis.";
 RL (1) Plant Gene Register PGR97-077.
 RN [2]

SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Nambu K., Okumura S., Shino S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Huang B., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoeneking T., Pepin K., Spieth J., Sekhon J., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latrelle P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Stromwall C.,
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,

RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vii D., Baker J.,
 RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
 RA Marijnsen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Voicaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
 RA Entian K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
 RA Langham S.A., McCullagh B., Roben J., Grymoprez B., Zimmermann W.,
 RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirse W., Mooijman P., Klein Lankhorst R.,
 RA Weitzenecker T., Bothé G., Rose M., Hauf J., Berner S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villarejo R., Gijlen J., Ardiles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K., Rüd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.,
 RA "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 thaliana".
 RL Nature 408:823-826(2000).
 RN [3]

CHARACTERIZATION.
 RP STRAIN=cv. C24;
 RX MEDLINE=21260187; PubMed=11359606;
 RA Ledger S., Strayer C., Ashton F., Kay S.A., Puterill J.;
 RT "Analysis of the function of two circadian-regulated CONSTANS-LIKE
 genes";
 RL Plant J. 26:15-22(2001).
 CC -1- FUNCTION: Putative transcription factor that may be involved in
 the light input to the circadian clock but does not affect
 flowering time.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LEAVES AND AT LOWER LEVELS
 IN STEMS, FLOWERS AND STILICES. NOT DETECTED IN ROOTS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT.
 CC -1- INDUCTION: EXPRESSED WITH A CIRCADIAN RHYTHM SHOWING A PEAK
 AT DAWN.
 CC -1- SIMILARITY: BELONGS TO THE CONSTANS FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 B BOX-TYPE ZINC FINGERS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: Y10555; CAA71587.1; -.
 DR EMBL: Y10556; CAA71588.1; -.
 DR EMBL: AL391144; CAC01784.1; -.
 DR InterPro: IPR000315; Znf_Box.
 DR InterPro: IPR002926; Znf_Constans.
 DR Pfam: PF00643; Zf-B_Box; 2.
 DR ProDom: PD007661; Znf_Constans; 2.
 DR SMART: SM00336; BBOX; 1.
 DR PROSITE: PS50119; ZF_BBOX; 2.
 KW Transcription regulation; Zinc-finger; DNA-binding; Nuclear protein;
 KW Repeat; Multigene family.
 FT ZN_FING 12 54 B BOX-TYPE 1.
 FT ZN_FING 55 97 B BOX-TYPE 2.
 SQ SEQUENCE 355 AA; 39491 MW; 412C237A9CD30309 CRC64;

Query Match 8.4%; Score 67.5; DB 1; Length 355;
 Best Local Similarity 23.6%; Pred. No. 11;
 Matches 29; Conservative 15; Mismatches 44; Indels 35; Gaps 4;

QY 20 PWANICAGKSSNEIRTCRRGCGYSQSRPHQGVDXL-----CSAGSTV 66
 DB 62 PAAPFCKADAACCTTCDSE---IHSANPLARRHQRVILPISEVSYSSTATNHSCETV 118
 QY 67 YAPFTGIVGQEK-----PYQKNAINNGVRISGRGCVKMFYIKPIKYKCP 113
 DB 119 TDPENLVIGQEBDEDEDAASWLLPNSGKSGNN-----NGFSIGDFLNLVYSSS 172
 QY 114 IKK 116
 DB 173 DRQ 175


```

FT SIGNAL      1 30
FT CHAIN      31 625
FT            31 625
FT DOMAIN     31 214
FT TRANSMEM   215 234
FT DOMAIN     235 625
FT REPEAT     35 69
FT REPEAT     72 113
FT REPEAT     115 152
FT REPEAT     155 195
FT DISULFID   35 47
FT DISULFID   48 61
FT DISULFID   51 69
FT DISULFID   72 87
FT DISULFID   93 113
FT DISULFID  115 128
FT DISULFID  134 152
FT CARBOHYD  106 106
FT CARBOHYD  175 175
FT CONFLICT  494 494
SQ SEQUENCE  625 AA; 66621 MW; F8C1872E99511D8E CRC64;

Query Match      8.48; Score 67.5; DB 1; Length 625;
Best Local Similarity 21.48; Pred. No. 20;
Matches 31; Conservative 11; Mismatches 40; Indels 63; Gaps 5;

QY 16 ALAGPWANI-----CAGKSSNEIRTCDRHGCGQYSAQRSQRPHQGVDXLCSA 62
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 309 AAGGPWAEVRDSRTFTLVSEVETQGDLSRKIPTDEY-----TDRPSQPTGSLLIQQ 362
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 63 GSTVYAPFTGMIVGQEKPYQKNKAINNGVRISGRGFCVKMFYIKPKYKGPVKKGLGT 122
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 363 GSKSIPPF-----QE-----PLEVGENDS- 381
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 123 LLPLQKVPYPGIOSHVHIENCDSSDP 147
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 382 ---LSQCTGTSTVDSEGCDFTEP 403
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

```

Search completed: March 25, 2003, 17:20:05
Job time : 16.5 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:16:20 ; Search time 28.5 seconds

(without alignments)
1091.688 Million cell updates/sec

Title: US-09-868-953A-1

Perfect score: 804

Sequence: 1 MFSTKALLAGLISTALGP.....GTQSHVHTNCSSDPTAYL 151

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp-archaea:*
- 2: sp-bacteria:*
- 3: sp-fungi:*
- 4: sp-human:*
- 5: sp-invertebrate:*
- 6: sp-mammal:*
- 7: sp-mhc:*
- 8: sp-organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-virus:*
- 13: sp-vertebrate:*
- 14: sp-unclassified:*
- 15: sp-rvirus:*
- 16: sp-bacteriap:*
- 17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Description |
|------------|-------|-------------|--------------------|
| 1 | 418.5 | 52.1 | Q9dge4 cyprinus ca |
| 2 | 370 | 46.0 | Q9dfj1 oncorhynch |
| 3 | 142.5 | 17.7 | Q21241 caenorhabdi |
| 4 | 87.5 | 10.9 | Q01719 strongyloce |
| 5 | 82 | 10.2 | Q96ze8 sulfolobus |
| 6 | 80.5 | 10.0 | Q9pgx9 xylella fas |
| 7 | 78 | 9.7 | Q9ptr2 fundulus he |
| 8 | 77.5 | 9.6 | Q9ied5 human immun |
| 9 | 76 | 9.5 | Q9hs60 halobacteri |
| 10 | 76 | 9.5 | Q9uf93 homo sapien |
| 11 | 74.5 | 9.3 | Q98679 cricetus |
| 12 | 74 | 9.2 | Q98q88 mycoplasma |
| 13 | 73.5 | 9.1 | Q9xvj5 raistonia s |
| 14 | 73.5 | 9.1 | Q9cp23 pasteurella |
| 15 | 73 | 9.1 | Q91hq4 hepatitis c |
| 16 | 73 | 9.1 | Q8zah3 versinia pe |

| | | | | | | |
|----|------|-----|------|----|--------|--------------------|
| 17 | 73 | 9.1 | 877 | 15 | Q8Q715 | Q8Q715 human immun |
| 18 | 72.5 | 9.0 | 284 | 2 | Q9AKK4 | Q9AKK4 rickettsia |
| 19 | 72.5 | 9.0 | 900 | 15 | Q9QNZ8 | Q9QNZ8 human immun |
| 20 | 72.5 | 9.0 | 1509 | 10 | Q8S2E1 | Q8S2E1 oryza sativ |
| 21 | 72 | 9.0 | 478 | 4 | Q9BSK1 | Q9BSK1 homo sapien |
| 22 | 71.5 | 8.9 | 295 | 16 | Q8LYL3 | Q8LYL3 anabaena sp |
| 23 | 71.5 | 8.9 | 706 | 10 | Q81324 | Q81324 arabidopsis |
| 24 | 71.5 | 8.9 | 749 | 10 | Q9LLU8 | Q9LLU8 arabidopsis |
| 25 | 71.5 | 8.9 | 773 | 10 | Q9C902 | Q9C902 arabidopsis |
| 26 | 71 | 8.8 | 190 | 12 | Q9J3A1 | Q9J3A1 rickettsia |
| 27 | 71 | 8.8 | 284 | 16 | Q92GNO | Q92GNO rickettsia |
| 28 | 70.5 | 8.8 | 548 | 15 | Q9IED6 | Q9IED6 human immun |
| 29 | 70.5 | 8.8 | 871 | 15 | Q8Q7J2 | Q8Q7J2 human immun |
| 30 | 70 | 8.7 | 270 | 16 | Q8RD77 | Q8RD77 thermoaer |
| 31 | 70 | 8.7 | 352 | 2 | Q9AHR2 | Q9AHR2 uncultured |
| 32 | 70 | 8.7 | 444 | 10 | Q9SUP5 | Q9SUP5 arabidopsis |
| 33 | 69.5 | 8.6 | 644 | 17 | Q974G9 | Q974G9 sulfolobus |
| 34 | 69.5 | 8.6 | 263 | 11 | Q9WU52 | Q9WU52 mus musculu |
| 35 | 69.5 | 8.6 | 263 | 11 | Q9R1O6 | Q9R1O6 mus musculu |
| 36 | 69.5 | 8.6 | 268 | 16 | Q97JM2 | Q97JM2 clostridium |
| 37 | 69.5 | 8.6 | 715 | 10 | Q9LNL9 | Q9LNL9 arabidopsis |
| 38 | 69.5 | 8.6 | 873 | 15 | Q8Q7G4 | Q8Q7G4 human immun |
| 39 | 69 | 8.6 | 186 | 12 | Q9J3D9 | Q9J3D9 hepatitis c |
| 40 | 69 | 8.6 | 191 | 12 | Q9Q3X9 | Q9Q3X9 hepatitis c |
| 41 | 69 | 8.6 | 473 | 2 | Q8VUO1 | Q8VUO1 pseudomonas |
| 42 | 69 | 8.6 | 1262 | 5 | Q95YG2 | Q95YG2 caenorhabdi |
| 43 | 69 | 8.6 | 1319 | 5 | Q44518 | Q44518 caenorhabdi |
| 44 | 68.5 | 8.5 | 734 | 2 | Q85122 | Q85122 shewanella |
| 45 | 68.5 | 8.5 | 735 | 2 | Q87539 | Q87539 shewanella |

ALIGNMENTS

RESULT 1

Q9DGE4 PRELIMINARY; PRT; 160 AA.
AC Q9DGE4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myeloid protein-1.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP FUJIKI K., NAKAO M., SHIN D., YANO T.;
RA "Molecular cloning of a carp homolog of chicken myeloid protein-1."
RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB027192; BAB16024.1;
SQ SEQUENCE 160 AA; 17461 MW; BD15A9A10B4324C CRC64;

| | | | | | | | |
|-----------------------|-------|---|----------|------------|-----|--------|-------------|
| Query Match | 52.1% | Score | 418.5 | DB | 13 | Length | 160; |
| Best Local Similarity | 50.3% | Pred. No. | 8.3e-38; | | | | |
| Matches | 78; | Conservative | 28; | Mismatches | 36; | Indels | 13; Gaps 4; |
| QY | 6 | ALLAGLISTALA-----GPWANICAGKSSNEIRTC-D-RHGCGQYSAQR--SQRP HQG 55 | | | | | |
| DB | 8 | SFLLAVICSSLDVDSQVKGFP---LCSGSSNRKRKGCDDKKGCGNYGASRDGGRRHAG 64 | | | | | |
| QY | 56 | VDXLCSAGSTVYAPFTGMIVGQEKPYQNKNNAINNVRISGRFCVVMFYIKPKYKGP 115 | | | | | |
| DB | 65 | LDIVCADGATVYAPFDVFKLNGKAVPYKNNAINDINLSSGGLCFLFYVVKPISYFGTLK 124 | | | | | |
| QY | 116 | KGEKLTLLPLQKVPYQISQSHVHTENCSSDPTAY 150 | | | | | |
| DB | 125 | KGQKIGTMLPMQKVPYPGITSHVHVQMCDSRDP TKY 159 | | | | | |

RESULT 2

```
Q9DFJ1
ID Q9DFJ1 PRELIMINARY; PRT; 156 AA.
AC Q9DFJ1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Chemotaxin (Fragment).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Bayne C.J., Gerwick L., Fujiki K., Nakao M., Yano T.;
RT "Immune-relevant (including acute phase) genes identified in the
RT livers of rainbow trout, Oncorhynchus mykiss, by means of suppression
RT subtractive hybridization."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF271114; AAG28030.1;
FT NON_TER 156
SQ SEQUENCE 156 AA; 17087 MW; B19067665FDA70DA CRC64;

Query Match 46.0%; Score 370; DB 13; Length 156;
Best Local Similarity 46.6%; Pred. No. 1.6e-32;
Matches 68; Conservative 29; Mismatches 47; Indels 2; Gaps 1;

QY 7 LLAGLSTALAGPWANICAGKSSNEIRTCDRHGCGQYSAQRSPRHQGVDXLCSAGSTV 66
DB 10 VVLAVLSECEVMKFGQLCSDSSNRRTGDRWQGHGASRGRAHOGLLDIVNDGATV 69
QY 67 YAPFTGMIVGQEKPYQ--KNAINNGVRISGRGFCVKMFYIKPIKYGKPKIKKGLPTLL 124
DB 70 YAPFDVLNGKRVIVYDPKKAINDGINSGLGCLFELFVVKPKYSSVVKKGQRIGTML 129
QY 125 PLQKYPGIQSHVHIENCSDSDPTAY 150
DB 130 TMSQSVPGITSHRVQMCKSDPTKF 155

RESULT 3
Q21241
ID Q21241 PRELIMINARY; PRT; 472 AA.
AC Q21241;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 53.3 kDa protein.
GN K05F1.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Wohldmann P.;
RT "The sequence of C. elegans cosmid K05F1.5";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
```

```
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: U29377; AAA68720.2;
KW Hypothetical protein.
SQ SEQUENCE 472 AA; 53330 MW; D7A8D966A7190C2A CRC64;

Query Match 17.7%; Score 142.5; DB 5; Length 472;
Best Local Similarity 27.1%; Pred. No. 4.1e-07;
Matches 45; Conservative 27; Mismatches 63; Indels 31; Gaps 9;

QY 7 LLAGLSTALAGPWANICAGKSSNEIRTC--DRHG--CGQYSAQR--SORPHQGVDXLCSA 62
DB 146 LILIGIFIGGOTLGMKVCVGNNAKNDRCPTDGCSCNYHTERTSGEIDGVDVVRCHL 205
QY 63 GSTVIYAPFTGMIVGQEKPY---QNKAINNGVRISG-----RGFCVKMFYIKPIKYGKPIK 115
DB 206 GEPIYAPIEGEMFEP--RPGYGGKREKSCADQGVRIEGTGQWQGYAVHSSVKLSFFGGHVE 264
QY 116 KGEKLGTLPL---QKYPGIGTQSHV-----HIENC 142
DB 265 AGDEIGBALNRYCNDRGQNDVPEHVEIRLYKEGRLLDPTHHLQNC 310

RESULT 4
O01719
ID O01719 PRELIMINARY; PRT; 689 AA.
AC O01719;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Myb-related transcription factor (Fragment).
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98088684; PubMed=9428408;
RA Coffman J.A., Kirchhamer C.V., Harrington M.G., Davidson E.H.;
RT "SpMyb functions as an intramodular repressor to regulate spatial
RT expression of Cyllia in sea urchin embryos."
RL Development 124:4717-4727(1997).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 3 MYB-LIKE DOMAINS.
DR EMBL: U96090; AAC47807.1;
DR HSSP; P06876; IMBK.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00249; myb_DNA_binding; 3.
DR SMART; SM00395; SANT; 3.
DR PROSITE; PS00037; MYB_1; UNKNOWN_3.
DR PROSITE; PS00334; MYB_2; 3.
DR PROSITE; PS50090; MYB_3; 3.
KW DNA-binding; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 689 AA; 77241 MW; 0FA5A8D7AF8766C6 CRC64;

Query Match 10.9%; Score 87.5; DB 5; Length 689;
Best Local Similarity 22.4%; Pred. No. 0.65;
Matches 36; Conservative 22; Mismatches 46; Indels 57; Gaps 9;

QY 17 LAGPWANIC---AGKSSNEIR-----TCDR--HGCGQYSAQRSPR--PH---QGVDXILCS 61
DB 163 LGNRWAEIAKLLPGRTDNDNAIKHNWSTMRKRVETCNPTSKTKTTPHYTHTNDNQKPCS 222
QY 62 AGSTVIYAPFTGMIVGQEKPYQNKAINNGVRISGRGFCVKMFYIKPIKYGKPKIKGKGLG 121
DB 223 SSSKIYTP-----DSDFMNANSIRDALMRMGQGRV-----VR 255
QY 122 TLLPL-----QKYPGI---QSHVHIENCSDSDP 147
DB 256 TLYPMGHDTLQDEGEKSGKVGKTPQKWLIMDNCGEISP 296
```



```

RESULT 5
Q962E8 ID Q962E8 PRELIMINARY; PRT; 272 AA.
AC Q962E8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein S71885.
GN S71885.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Nakamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain".
RL DNA Res. 8:123-140(2001).
DR EMBL; AF000988; BAB66977.1; -.
DR InterPro; IPR000515; BPD_transp.
DR PROSITE; PS00402; BPD_TRANSF.INN.MEMBR; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 272 AA; 30871 MW; 810C97413773BF05 CRC64;

Query Match 10.2%; Score 82; DB 17; Length 272;
Best Local Similarity 29.7%; Pred. No. 0.88;
Matches 27; Conservative 18; Mismatches 34; Indels 12; Gaps 5;

QY 67 YAPFTGMI-----VQEKPYQNKNAINGVRGFCVKNFYIKPIKYKG-PIKKEG 118
Db 36 YSPFGIIIEKIEKVKIGKPNKY-TKIDVDVVMYINSNGRIKVLHVPYLEEGSEIKEGE 94
QY 119 KLTLLPLQKVY-PGIQSHVHIENCDSDPT 148
Db 95 KIGKF--LESFYTAGDFKHAIEGITFEPS 123

RESULT 6
Q9PGX9 ID Q9PGX9 PRELIMINARY; PRT; 417 AA.
AC Q9PGX9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein Xf0167.
GN Xf0167.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,

```

```

RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
DR EMBL; AE003870; AAF82980.1; -.
DR InterPro; IPR002886; Peptidase_M37.
DR Pfam; PF01551; Peptidase_M37; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 417 AA; 46554 MW; 5483743CD61061FD CRC64;

Query Match 10.0%; Score 80.5; DB 16; Length 417;
Best Local Similarity 25.0%; Pred. No. 2.1;
Matches 29; Conservative 24; Mismatches 42; Indels 21; Gaps 7;

QY 34 RTCDRHCGGQYSAORSQRPHQGVDXLCSAGSVYVAFPTGMI--VGQKPYQNKNAINGV 91
Db 249 RLSSTFGLRRPILGTMRMHKGVDAATGTPMAAGDARVQFQIGQYGVVILNH-- 306
QY 92 RISGRFCVKMFYIKPIKYKGIKKEKL--GFLPLQKVYPGIQS-----HVITE 140
Db 307 --GKGY--RTLYGHMSRF-GKIKAGQKINQGTIVG----YVGMTGLATGPHLHYE 352

RESULT 7
Q9PTR2 ID Q9PTR2 PRELIMINARY; PRT; 498 AA.
AC Q9PTR2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Cytochrome P450 2P2.
GN CYP2P2.
OS Fundulus heteroclitus (Killifish) (Mummichog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Fundulidae; Fundulus.
OX NCBI_TaxID=8078;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Oleksiak M.F., Stegeman J.J.;
RT "Diversity of vertebrate cytochrome P450 2 family genes: Cloning of
RT three new genes and phylogenetic analysis of Cyp2 gene subfamilies in
RT Fundulus heteroclitus";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Oleksiak M.F., Wu S., Parker C., Zeldin D.C., Stegeman J.J.;
RT "Cloning and expression of a novel teleost cytochrome P450, CYP2P3:
RT conservation of arachidonic acid epoxidase/19-Hydroxylase.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF117342; AAF21999.1; -.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.

```


QY 89 NGVIRSGRCVKMFYIKPIYKGPPIKKGKIGTLPLQKVVPGIQSHVHIE---NCDSS 145
 DB 232 NGLK-TRYAHASKVF---VKVGDIVKAGORTAL---IGRTGRATGPHLFEVHVNDVPQ 283
 QY 146 DPTAYL 151
 DB 284 NPVAF 289
 RESULT 14
 Q9CP23 PRELIMINARY; PRT; 531 AA.
 AC Q9CP23;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Hypothetical protein PM0243.
 GN PM0243.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida PM70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL; AE006058; AAK02327.1; .
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR InterPro; IPR002886; Peptidase_M37.
 DR Pfam; PF01551; Peptidase_M37; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 531 AA; 59590 MW; 93617D425C09F47D CRC64;

Query Match 9.1%; Score 73.5; DB 16; Length 531;
 Best Local Similarity 23.4%; Pred. No. 16;
 Matches 30; Conservative 18; Mismatches 47; Indels 33; Gaps 4;
 QY 24 ICAGKSSNEIRTCDRH-----GCGQYSQRSOR-----PHQ 54
 DB 346 ISAGKSYAIQASNGRYNRQGETLKGFPYLPQARVSSHFNYPYRHPVTGRIAPHK 405
 QY 55 GVDLCSAGSVYAPFTGMIVGOEKPYQNKAINNGVRISGRGFCVKMFYIKP--IKYKG 112
 DB 406 GVDAMPVGTPIAPADGIV--EKVAYQANGAGRYIVVRHGREYQTVYMLSLRALVKAGQ 463
 QY 113 PIKKGEL 120
 DB 464 NVKRGQRI 471

RESULT 15
 Q91HQ4 PRELIMINARY; PRT; 335 AA.
 ID Q91HQ4;
 AC Q91HQ4;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).
 DE Hepatitis C virus.
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Qian C., Zhi-Meng L., Ming Y.;
 RT "Variant Analysis and Immunogenecity Prediction of Envelope Gene of HCV Strains from China.";

RL Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.
 CC -!- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
 DR EMBL; AF403236; AAK95635.1; .
 DR InterPro; IPR002519; HCV_env.
 DR InterPro; IPR002531; HCV_NSI.
 DR Pfam; PF01539; HCV_env; 1.
 DR Pfam; PF01560; HCV_NSI; 1.
 DR ProDom; PD186062; HCV_NSI; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 1 335
 FT NON_TER 335 335
 SQ SEQUENCE 335 AA; 36215 MW; A3155D5CE2CDEDED1 CRC64;
 Query Match 9.1%; Score 73; DB 12; Length 335;
 Best Local Similarity 25.8%; Pred. No. 11;
 Matches 39; Conservative 19; Mismatches 65; Indels 28; Gaps 8;
 QY 8 LLAGLIALSTAGPWANIC-----AGKSSNEIRTCDRHCGGQYSQRSQ-RPHQGVDXLC 60
 DB 164 VLAGLAYYSMVGNWAKVLIVMLLFAGVDG-----GTYTTGGAQGRATQGLTSLF 212
 QY 61 SAGSTVYAPFTGMIVGOEKPYQNKAINNGVRISGRGFCVKMFYIKPIKYKG-PIKKGK 119
 DB 213 SRGS---AQKIQLVNTNGSWHINRTALNCNDSLQ-TGFLAALFYTHRFNSSGCP----ER 264
 QY 120 LGTLLPLQKVPYGIQSHVHIENCDSSDPTAY 150
 DB 265 LASCRPIDKFAQG-WGPITYDAPDSSDOKPY 294

Search completed: March 25, 2003, 17:21:07
 Job time : 33.5 secs